

Universidade Federal da Paraíba
Centro de Ciências e Tecnologia
Coordenação dos Cursos de Pós-Graduação em Engenharia Elétrica

Abordagem Híbrida para a Avaliação da Usabilidade de Interfaces com o Usuário

José Eustáquio Rangel de Queiroz

Tese submetida à Coordenação de Pós-Graduação em Engenharia Elétrica da Universidade Federal da Paraíba, como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências no Domínio da Engenharia Elétrica.

Área de Concentração: *Processamento da Informação*

M^a de Fátima Queiroz Vieira Turnell, PhD
Orientadora

Campina Grande
Junho de 2001



Q3a Queiroz, Jose Eustaquio Rangel de
Abordagem hibrida para a avaliacao da usabilidade de
interfaces com o usuario / Jose Eustaquio Rangel de
Queiroz. - Campina Grande, 2001.
410 p.

Tese (Doutorado em Engenharia Eletrica) - Universidade
Federal da Paraiba, Centro de Ciencias e Tecnologia.

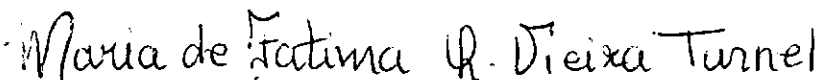
1. Interacao Homem-Maquina 2. Avaliacao de Produtos de
Software 3. Usabilidade de Produtos 4. Tese - Engenharia
Eletrica I. Turnell, Maria de Fatima Queiroz Vieira II.
Universidade Federal da Paraiba - Campina Grande (PB) III.
Titulo

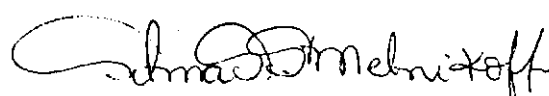
CDU 621.3:004.5(043)

**ABORDAGEM HÍBRIDA PARA A AVALIAÇÃO DA USABILIDADE DE
INTERFACES COM O USUÁRIO**

JOSÉ EUSTÁQUIO RANGEL DE QUEIROZ


Tese Aprovada em 26.06.2001


PROF. MARIA DE FÁTIMA QUEIROZ VIEIRA TURNELL, Ph.D., UFPB
Orientadora


PROFa. SELMA SHIN SHIMIZU MELNIKOFF, Dr., USP-SP
Componente da Banca


PROF. FERNANDO DA FONSECA DE SOUZA, Ph.D., UFPE
Componente da Banca


PROF. ANGELO PERKUSICH, D.Sc., UFPB
Componente da Banca


PROF. BERNARDO LULA JÚNIOR, Dr., UFPB
Componente da Banca

CAMPINA GRANDE - PB
Junho - 2001

**Às minhas mães, Nêga e Lourdes, *in*
*memoriam***

Ao meu pai, Inácio Queiroz

AGRADECIMENTOS

Em geral, cada etapa encerrada na vida acarreta dois sentimentos conflitantes: júbilo e perda. Júbilo pela finalização da tarefa, pela sensação do dever cumprido. Perda pelo distanciamento do contexto em que a tarefa se desenrolou e, sobretudo, pelo afastamento inevitável daqueles que contribuíram, cada um ao seu modo, para que a tarefa fosse cumprida.

Os agradecimentos são tantos que a memória pode falhar na tentativa de enumerá-los:

A minha família, pelo apoio cotidiano e pelo estímulo constante, os quais me impulsionaram sempre a prosseguir.

À Prof^a. M^a de Fátima Turnell, pela orientação séria e criteriosa, pelas recomendações sempre pertinentes e, acima de tudo, pela amizade inestimável.

Ao Prof. Antonio Marcus Nogueira Lima, pelo apoio valioso no momento preciso, assim como pelas discussões sobre o *MATLAB*.

Ao CNPq e à CAPES, sem o apoio dos quais a conclusão deste trabalho não teria sido possível.

Aos membros da banca examinadora, pela gentileza em participar deste processo, oferecendo recomendações e sugestões para a otimização da qualidade deste documento.

Ao Prof. Juarez Fernandes, pela discussão esclarecedora sobre o processamento estatístico dos dados, assim como pela revisão posterior dos resultados dos procedimentos aplicados.

A todos os participantes do ensaio de usabilidade, pela disponibilidade, atenção e seriedade com que participaram da pesquisa.

A todos os amigos do LAPS e do LMRS/PB, em especial à amiga Joseana Macedo, pelo convívio amigável e pelas sugestões oferecidas ao longo da realização da pesquisa.

Ao amigo Fernando Araújo, pela ajuda significativa na implementação dos instrumentos de sondagem no formato HTML.

A todos aqueles que, direta ou indiretamente, ofereceram sua contribuição para a concretização deste trabalho.

A Deus, fonte inesgotável de sabedoria, minha fortaleza em todos os instantes da vida e minha luz nas horas sombrias.

RESUMO

Neste trabalho são confrontadas três categorias de enfoques de avaliação da usabilidade de produtos de *software*, a saber: *mensuração do desempenho do usuário durante o uso do produto*, *inspeção de conformidade do produto a um padrão internacional* e *sondagem da satisfação do usuário no tocante ao produto*. O propósito do trabalho foi investigar a influência dos graus de satisfação, desempenho e conformação a um padrão internacional sobre a usabilidade do produto. A pesquisa consistiu na investigação da relação entre os pares de fatores de avaliação da usabilidade: (i) desempenho e satisfação do usuário; (ii) desempenho do usuário e conformação do produto a um padrão; e (iii) satisfação do usuário e conformação do produto a um padrão.

A partir de um conjunto de sete hipóteses, foi desenvolvido um estudo analítico, objetivando testá-las. Foram realizadas sessões de inspeção do produto com base nas *Partes 14, 15 e 16* do padrão internacional *ISO 9241*. Posteriormente, um ensaio de usabilidade foi planejado e conduzido em ambiente controlado (laboratório de usabilidade), tendo contado com a participação de 40 usuários de teste. Foram coletados dados de natureza quantitativa e qualitativa: (i) perfis e opiniões dos usuários de teste sobre a usabilidade do produto; (ii) registros escritos e em vídeo do desempenho do usuário durante o uso do produto; e (iii) julgamentos de avaliadores, fundamentados no padrão *ISO 9241*. Sobre os dados tabulados, foram realizadas análises estatísticas e não estatísticas. Por fim, foram formuladas conclusões, embasadas pelos resultados dos testes estatísticos de hipóteses.

Os resultados mostram que o enfoque avaliatório adotado produziu um quadro mais completo das falhas de usabilidade do produto do que a aplicação isolada de qualquer uma das técnicas consideradas. A pesquisa também resultou em uma descrição detalhada dos três enfoques investigados, com recomendações sobre *quando* e *como* aplicá-los, individualmente ou integrados em uma abordagem híbrida.

A seleção de métodos de avaliação da usabilidade é usualmente heurística, fundamentada na experiência prévia do avaliador. Adicionalmente, a literatura disponível normalmente não detalha suficientemente os métodos de avaliação propostos. Este trabalho discute não somente uma série de critérios a serem considerados durante a seleção de métodos, como também oferece uma base sólida para suportar o processo de seleção.

Os enfoques investigados produziram resultados complementares no tocante à identificação de falhas na usabilidade do produto avaliado. Apesar dos resultados obtidos a partir dos diferentes enfoques adotados terem se mostrado consistentes entre si, foram registradas discrepâncias que sugerem a combinação de enfoques avaliatórios como meio de se obter uma visão mais abrangente das questões de usabilidade de produtos.

ABSTRACT

This work confronts three major strategies used in software product usability evaluation: *usability testing through the observation of the product in use*, *usability inspection based upon international standards* and, *user satisfaction survey*. The purpose is to investigate the influence of levels of satisfaction, levels of performance and levels of compliance (*adherence ratings*) on the product's usability. The research consisted on investigating the relationship between the pairs of factors: (i) user performance and satisfaction, (ii) user performance and compliance to a standard, and (iii) user satisfaction and compliance to a standard.

A set of seven hypotheses was formulated, followed by an analytical study in order to investigate them. Experiments were then planned and performed with a set of 40 users, resulting in a rich set of data. The data consisted of (i) information on user profiles and user opinions about the product usability, (ii) paper and video observations of the user performance during product use in a controlled environment (usability laboratory); and (iii) product inspection checklists with the international standard ISO 9241. The data were then tabulated and statistically analysed. Finally the findings were synthesised and conclusions were drawn on the results yielded by the hypothesis tests.

The results point towards a hybrid evaluation method in order to obtain a clear picture of the product usability flaws and strengths. The work also resulted in a detailed description of all three investigated strategies with recommendations on *when* and *how* to apply them, solely or combined in a hybrid approach. In order to support the decision making on choosing a method or a combination of methods, the results were tabulated to facilitate their use.

The choice of usability evaluation methods is usually heuristic, based upon the evaluator's previous experience. The available literature does not make clear the detail behind the application of a given method. This work not only discusses a set of parameters to be considered during the selection of methods, but it also gives a solid basis to support this choice.

The investigated methods complement each other in their findings about product usability. The results are consistent between the methods but there are differences between them that suggest a combination of methods in order to obtain a complete picture of the usability questions.

SUMÁRIO

1	Considerações Iniciais	1
1.1	Introdução	3
1.2	Caracterização do Objeto da Pesquisa	7
1.3	Motivação e Justificativa da Pesquisa	14
1.4	Objeto da Pesquisa	17
1.5	Objetivos da Pesquisa	19
1.5.1	Geral.....	19
1.5.2	Específicos	19
1.6	Encaminhamento Metodológico.....	19
1.7	Organização do Documento	21
2	Considerações sobre Qualidade e Usabilidade de Software	23
2.1	Qualidade de Software e Satisfação do Usuário	25
2.1.1	Qualidade e Qualidade de Software	26
2.3.2	Facetas da Qualidade de Software	34
2.3.2	Qualidade de Software e Satisfação do Usuário	39
2.2	Engenharia, Engenharia de Software e Engenharia da Usabilidade.....	41
2.2.1	Engenharia.....	41
2.2.2	Engenharia de Software.....	43
2.2.3	Engenharia da Usabilidade.....	45
2.3	Usabilidade de Software	47
2.3.1	Visão de Shackel.....	53
2.3.2	Visão de Nielsen	55
2.3.3	Visão da ISO (<i>International Standards Organization</i>).....	57
3	Métodos de Avaliação de Interfaces Usuário-Computador	63
3.1	Métodos de Avaliação de Interfaces Usuário-Computador.....	65
3.2	Métodos de Avaliação da Usabilidade	68
3.2.1	Ensaio de Usabilidade (<i>Usability Testing</i>).....	69
3.2.1.1	Observações (<i>Observation</i>).....	72
3.2.1.2	Uso de Questionários (<i>Questionnaires</i>).....	75
3.2.1.3	Entrevistas (<i>Interviews</i>).....	77

3.2.1.4	Verbalização de Procedimentos (<i>Thinking Aloud</i>)	79
3.2.1.5	Interação Construtiva (<i>Constructive Interaction</i>)	80
3.2.1.6	Ensaio Retrospectivo (<i>Retrospective Testing</i>)	81
3.2.1.7	Captura Automática a partir da Aplicação (<i>Automatic Logging from the Application</i>)	82
3.2.1.8	Discussões em Grupo (<i>Focus Groups</i>)	84
3.2.1.9	Retorno de Opiniões do Usuário (<i>User Feedback</i>) / Ensaio de Usabilidade Remoto (<i>Remote Usability Testing</i>).....	86
3.2.2	Inspeções de Usabilidade (<i>Usability Inspections</i>).....	91
3.2.2.1	Revisões Sistemáticas (<i>Walkthroughs</i>).....	91
3.2.2.2	Diretrizes de Projeto, Guias de Estilo e Padrões (<i>Design Guidelines, Styleguides and Standards</i>)	99
3.2.2.3	Avaliação Heurística (<i>Heuristic Evaluation</i>).....	108
3.2.2.4	Inspeção Fundamentada na Perspectiva (<i>Perspective-based Inspection</i>).....	114
3.3	Considerações Finais	115
4	Enfoques Avaliatórios da Usabilidade de Interfaces Usuário-Computador.....	118
4.1	Introdução	119
4.2	Enfoques Avaliatórios Consagrados	120
4.3	Mensuração da Usabilidade	123
4.3.1	Enfoques Avaliatórios e Pontos de Vista de Processos Interativos.....	124
4.3.2	Seleção do Enfoque Avaliatório.....	127
4.4	Caráter Complementar dos Enfoques Quantitativo e Qualitativo na Pesquisa.....	128
5	Inspeção de Conformidade de Produtos com o Padrão ISO 9241	137
5.1	Importância e Benefícios Advindos da Padronização.....	139
5.2	Aspectos da Padronização da Interação Homem-Máquina.....	145
5.3	Padrões Destinados à Interação Homem-Máquina	148
5.4	O Padrão <i>ISO 9241</i>	153
5.5	Metodologia Adotada na Inspeção do Produto-Alvo com Base nas <i>Partes 11, 14, 15 e 16 do Padrão ISO 9241</i>	158
5.5.1	Aplicação das <i>Partes 14, 15 e 16 do Padrão ISO 9241</i>	160
5.5.1.1	Adequação de Diálogos via Menus, Linguagens de Comandos e Manipulação Direta	160
5.5.1.2	Aplicabilidade das Recomendações da ISO.....	162
5.5.1.3	Síntese dos Aspectos Interativos Abordados pelas <i>Partes 14, 15 e 16 do Padrão ISO 9241</i>	163
5.5.2	Considerações sobre a Aplicabilidade e Adoção das <i>Partes 14, 15 e 16 do Padrão ISO 9241</i>	164
5.5.2.1	Avaliação de Produtos.....	164
5.5.2.2	Aplicabilidade	165
5.5.2.3	Métodos para a Determinação da Aplicabilidade e/ou Adoção das Recomendações	166

5.5.3	Uso das Listas de Inspeção (<i>Checklists</i>) das Partes 14, 15 e 16.....	169
5.5.4	Sumário dos Resultados de uma Inspeção.....	171
5.6	Comentários Finais.....	172
6	Sondagem da Satisfação Subjetiva e Mensuração do Desempenho do Usuário	173
6.1	Importância e Benefícios da Sondagem da Satisfação Subjetiva do Usuário	175
6.2	Importância e Benefícios da Mensuração do Desempenho do Usuário.....	176
6.3	Sondagem da Satisfação Subjetiva do Usuário.....	178
6.3.1	Questionários Disponíveis no Mercado para a Sondagem da Satisfação Subjetiva do Usuário de Sistemas Computacionais	178
6.3.1.1	QUIS (<i>Questionnaire for User Interaction Satisfaction</i>).....	179
6.3.1.2	ASQ (<i>After-Scenario Questionnaire</i>), PSSUQ (<i>Post-Study System Usability Questionnaire</i>) e CSUQ (<i>Computer System Usability Questionnaire</i>)	180
6.3.1.3	SUS (<i>System Usability Scale</i>).....	182
6.3.1.4	IsoMetrics	184
6.3.1.5	SUMI (<i>Software Usability Measurement Inventory</i>).....	186
6.3.2	Instrumentos de Sondagem Desenvolvidos no Âmbito desta Pesquisa.....	188
6.3.2.1	DePerUSI (<i>Definimento do Perfil do Usuário de Sistemas Interativos</i>).....	192
6.3.2.2	OpUS (<i>Opinião do Usuário de Software</i>)	195
6.4	Metodologia Adotada.....	197
6.4.1	Planejamento do Ensaio Avaliatório	198
6.4.1.1	Caracterização do Alvo de Estudo	198
6.4.1.2	Definição das Metas e Interesses	199
6.4.1.3	Caracterização do Universo Amostral	199
6.4.1.4	Levantamento dos Usuários de Teste Potenciais	199
6.4.1.5	Definição do Modo de Recrutamento dos Participantes.....	199
6.4.1.6	Decisão do Número de Participantes do Ensaio Avaliatório	200
6.4.1.7	Seleção das Técnicas de Avaliação	200
6.4.1.8	Definição de Indicadores Objetivos e Subjetivos.....	201
6.4.2	Elaboração do Material do Ensaio.....	204
6.4.2.1	Planejamento e Estruturação de Tarefas de Teste.....	204
6.4.2.2	Elaboração da Ficha Cadastral do Participante e do Documento de Aceitação das Condições de Teste	208
6.4.2.3	Elaboração do Material Necessário à Condução do Processo Avaliatório	208
6.4.2.4	Discussão das Abordagens de Avaliação da Usabilidade do Produto	208
6.4.2.5	Validação do Material Elaborado	208
6.4.3	Condução do Ensaio e Coleta de Dados	209
6.4.4	Tabulação e Análise dos Dados	209
6.4.4.1	Triagem Preliminar dos Dados Coletados.....	209
6.4.4.2	Triangulação dos Dados Coletados	210
6.4.4.3	Tabulação e Síntese dos Dados Coletados.....	210
6.4.4.4	Organização dos Problemas Listados	210
6.4.5	Apresentação dos Resultados	210

7	Apresentação e Discussão dos Resultados	211
7.1	Tipos de Dados Obtidos	213
7.2	Resultados da Inspeção de Conformidade do Produto com o Padrão Internacional <i>ISO 9241</i>	214
7.2.1	Problemas Identificados a partir da Inspeção da Conformidade do Produto à <i>ISO 9241-14</i>	216
7.2.2	Problemas Identificados a partir da Inspeção da Conformidade do Produto à <i>ISO 9241-15</i>	221
7.2.3	Problemas Identificados a partir da Inspeção da Conformidade do Produto à <i>ISO 9241-16</i>	228
7.2.3	Taxas de Adoção e Parecer sobre o Produto com base no Processo de Inspeção	236
7.3	Resultados da Mensuração do Desempenho	239
7.3.1	Análise Preliminar dos Indicadores Quantitativos	239
7.3.2	Análise Estatística dos Indicadores Quantitativos	249
7.3.3	Problemas Identificados a partir da Mensuração do Desempenho do Usuário	258
7.3.4	Parecer sobre o Produto com base na Mensuração do Desempenho do Usuário	268
7.4	Resultados da Sondagem da Opinião do Usuário	271
7.4.1	Resultados da Análise do DePerUSI	271
7.4.2	Resultados da Análise do OpUS	284
7.4.2.1	Análise Preliminar dos Escores Obtidos	289
7.4.2.2	Análise Estatística dos Indicadores Quantitativos	291
7.4.2.3	Problemas Identificados a partir da Sondagem da Satisfação do Usuário	294
7.5	Síntese dos Resultados Apresentados	298
8	Considerações Finais	303
8.1	Conclusões	304
8.2	Aplicabilidade dos Resultados da Pesquisa	309
8.3	Proposições para Trabalhos Futuros	311
	Referências Bibliográficas	313

LISTA DE FIGURAS

Fig. 1 – Qualidade e valores, segundo a <i>European Organisation for Quality</i> [EOQ00]	31
Fig. 2 – Fatores de qualidade de <i>software</i> definidos por McCall et al. [McCa77]	34
Fig. 3 – Critérios e fatores do modelo <i>SQM</i> [McCa77]	35
Fig. 4 – Características e sub-características do modelo contido no padrão <i>ISO 9126</i> [ISO92]	36
Fig. 5 – Critérios e fatores do modelo <i>Sintese do SQM</i> (ou <i>GRCM</i>) [Terv96] e comparação dos critérios do <i>GRCM</i> com as características do padrão <i>ISO 9126</i> [ISO92]	38
Fig. 6 – Aceitação do produto, conforme Shackel [Shac91], com destaque para as dimensões e medidas concretas da <i>usabilidade</i>	54
Fig. 7 – Concepção de Nielsen [Niel93c] para a <i>aceitabilidade</i> de um sistema, com destaque para as dimensões da <i>usabilidade</i>	56
Fig. 8 – Concepção de <i>usabilidade</i> da ISO, através do padrão <i>ISO9241-11</i> [ISO98]	58
Fig. 9 – Dimensões da <i>usabilidade</i> de acordo com a Parte 11 do <i>ISO 9241</i> [ISO98]	58
Fig. 10 – Agrupamento de <i>fatores de projeto</i> , segundo Treu [Treu94]	129
Fig. 11 – Exemplos de interpretação de números de referência de documentos ISO ...	152
Fig. 12 – Formato das recomendações individuais do padrão <i>ISO 9241</i>	162
Fig. 13 – Processo de decisão – Situação de avaliação (<i>Parte 14, 15 ou 16</i> do <i>ISO 9241</i>)	168
Fig. 14 – Excerto da lista de inspeção contida no <i>Anexo A</i> das <i>Partes 14, 15 e 16</i> do <i>ISO 9241</i>	170

Fig. 15 – Excertos do QUIS (<i>Questionnaire for User Interaction Satisfaction</i>): (a) versão impressa; e (b) versão HTML	180
Fig. 16 – Estrutura do ASQ (<i>After-Scenario Questionnaire</i>)	181
Fig. 17 – Excerto do PSSUQ (<i>Post-Study System Usability Questionnaire</i>)	181
Fig. 18 – Excerto do CSUQ (<i>Computer System Usability Questionnaire</i>)	182
Fig. 19 – Estrutura do SUS (<i>System Usability Scale</i>)	183
Fig. 20 – Excerto original da versão longa do <i>IsoMetrics 2.01</i>	185
Fig. 21 – Excerto original da versão curta do <i>IsoMetrics 2.01</i>	186
Fig. 22 – Excertos do SUMI (<i>Software Usability Measurement Inventory</i>): (a) instrumento de delineamento do perfil do usuário; e (b) quatro itens iniciais do instrumento da sondagem da opinião do usuário	187
Fig. 23 – Excerto do cabeçalho e declarações da primeira seção (<i>O SISTEMA COMO UM TODO</i>) de um dos precursores dos instrumentos <i>DePerUSI</i> e <i>OpUS</i>	189
Fig. 24 – Excerto seção referente ao delineamento do perfil técnico do respondente (<i>VOCÊ E O SITIM</i>) de um dos precursores dos instrumentos <i>DePerUSI</i> e <i>OpUS</i>	190
Fig. 25 – Excerto do <i>questionário pré-teste</i> , uma reestruturação do questionário elaborado por Queiroz [Quei94], feita por Queiroz e Turnell [Quei98a, Quei98b]	191
Fig. 26 – Excerto do <i>questionário pós-teste</i> , uma reestruturação do questionário elaborado por Queiroz [Quei94], feita por Queiroz e Turnell [Quei98a, Quei98b]	192
Fig. 27 – Excertos do <i>DePerUSI</i> : (a) versão impressa; e (b) versão <i>.xls</i>	194
Fig. 28 – Excerto da seção inicial do instrumento desenvolvido para a sondagem da <i>Opinião do Usuário de Software</i> (versão <i>OpUS.xls</i>)	196
Fig. 29 – Excerto da seção final do instrumento desenvolvido para a sondagem da <i>Opinião do Usuário de Software</i> (versão <i>OpUS.xls</i>)	197

Fig. 30 – Vistas aéreas do ambiente de condução do ensaio de usabilidade (LIHM/DEE/CCT/UFPB) do ponto de vista da: (a) entrada; e (b) face posterior	206
Fig. 31 – Cronograma de condução do ensaio	207
Fig. 32 – Técnica gráfica utilizada pela interface do <i>Windows Explorer</i> para indicação das ações do usuário no processo de seleção de itens de interesse da Ajuda	216
Fig. 33 – Técnica gráfica utilizada pela interface do <i>MATLAB v. 5.3.0</i> para a indicação das ações do usuário no processo de seleção de itens de interesse da Ajuda	217
Fig. 34 – Destaque para a opção <i>Window</i> do menu principal do <i>MATLAB</i> e suas sub-opções	219
Fig. 35 – Destaque para a falha no comportamento cíclico do menu principal do <i>MATLAB</i>	219
Fig. 36 – Mecanismos de interação do usuário com as facilidades oferecidas pelo <i>MATLAB v. 5.3.0</i> , com destaque para a falha referente à recomendação 5.3.1 (Geral)	223
Fig. 37 – Destaque para a falha do <i>MATLAB v. 5.3.0</i> no tocante à recomendação 6.4 (Correção de erros) da <i>ISO 9241-15</i>	224
Fig. 38 – Destaque para a falha do <i>MATLAB v. 5.3.0</i> referente à correção de erros (recomendação 6.4 da <i>ISO 9241-15</i>)	225
Fig. 39 – Destaque para a falha do <i>MATLAB v. 5.3.0</i> referente a comandos destrutivos	226
Fig. 40 – Destaque para a falha no comportamento do botão <i>Home</i>	229
Fig. 41 – Destaque para a falha no comportamento do botão <i>Tips</i>	229
Fig. 42 – Destaque para a falha da interface do <i>Editor/Debugger</i> da versão 5.3.0 do <i>MATLAB</i> relativa à recomendação 6.1.5 da <i>ISO 9241-16</i>	230
Fig. 43 – Destaque para a falha da interface do <i>Simulink</i> relativa ao arrasto de objetos	231

Fig. 44 – Redimensionamento de objetos gráficos no <i>MS Word2000</i> e no <i>MATLAB v. 5.3.0</i>	233
Fig. 45 – Janela de um modelo (arquivo <i>.mdl</i>) e janelas de resultados do <i>Simulink</i>	234
Fig. 46 – Detalhe da ampliação de um resultado gráfico em uma janela de resultados do <i>Simulink</i>	234
Fig. 47 – Barra de ferramentas da janela de resultados do <i>Simulink</i>	235
Fig. 48 – Falha do <i>MATLAB v. 5.3.0</i> relativa ao destaque e à clareza das mensagens de erros	259
Fig. 49 – Falha do <i>MATLAB v. 5.3.0</i> devido à inconsistência da similaridade de nomes e da sintaxe de frases de comando entre o <i>MS-DOS</i> e o <i>MATLAB v. 5.3.0</i>	260
Fig. 50 – Falha do <i>MATLAB v. 5.3.0</i> relativa à alteração das propriedades de objetos gráficos ou de texto na janela de figuras	262
Fig. 51 – Falha do <i>MATLAB v. 5.3.0</i> relativa ao retrocesso de ações (<i>undo</i>) após a alteração das propriedades de objetos gráficos ou de texto na janela de figuras	263
Fig. 52 – Barra de ferramentas de manipulação direta da janela de figuras do <i>MATLAB v. 5.3.0</i>	263
Fig. 53 – Falha do <i>MATLAB v. 5.3.0</i> relativa indicação visual da ativação do recurso de rotação tridimensional do sistema de eixos da janela de figuras	264
Fig. 54 – Estratégias de indicação visual nos processos de redimensionamento e rotação tridimensional de objetos no (a) <i>MS Word2000</i> e (b) <i>CorelDRAW 9</i>	265
Fig. 55 – Detalhamento do processo de rotação do plano do gráfico da janela de figuras do <i>MATLAB v. 5.3.0</i>	266
Fig. 56 – Destaque para a falha de projeto do painel de sub-opções da opção <i>Edit</i>	268
Fig. 57 – Distribuição numérica da <i>Formação Acadêmica</i> dos respondentes	273
Fig. 58 – Distribuição numérica do <i>Sexo</i> dos respondentes	274
Fig. 59 – Distribuição numérica da <i>Destreza Manual</i> dos respondentes	274

Fig. 60 – Distribuição numérica do <i>Uso de Corretivos Visuais</i> dentre os respondentes	275
Fig. 61 – Distribuição numérica da <i>Faixa Etária</i> dos respondentes	275
Fig. 62 – Distribuição numérica da <i>Natureza de Uso do Produto</i> dentre os respondentes	276
Fig. 63 – Distribuição numérica das <i>Atividades Desenvolvidas com o Auxílio do Produto</i> pelos respondentes	277
Fig. 64 – Distribuição numérica do <i>Tempo de Uso de Sistemas Computacionais</i>	277
Fig. 65 – Distribuição numérica da <i>Freqüência de Uso de Sistemas Computacionais</i>	278
Fig. 66 – Distribuição numérica do <i>Tempo de Uso do Produto</i>	279
Fig. 67 – Distribuição numérica do <i>Freqüência de Uso do Produto</i>	279
Fig. 68 – Distribuição numérica do aspecto <i>Versão Utilizada do MATLAB</i>	280
Fig. 69 – Distribuição numérica da <i>Forma de AJUDA mais Utilizada</i>, segundo os respondentes	281
Fig. 70 – Distribuição numérica das <i>Caixas de Ferramentas (Toolboxes)</i>	281
Fig. 71 – Distribuição numérica da <i>Experiência Prévia com Produtos Similares</i> dos respondentes	282
Fig. 72 – Distribuição numérica dos produtos similares, segundo os respondentes	282
Fig. 73 – Distribuição numérica do <i>Tempo de Uso de UM Produto Similar ao MATLAB</i>	283

Quadro 16 – Organizações Internacionais de Padronização	142
Quadro 17 – Corpos de Padronização Regional reconhecidos pela ISO, IEC e ITU no documento do padrão <i>ISO 9126 [ISO92]</i>	143
Quadro 18 – <i>Comissões Técnicas (CT)</i> do <i>SC-21:10</i> e suas <i>Comissões de Estudo (CE)</i>	144
Quadro 19 – Prefixos indicadores de tipos de publicações da ISO	149
Quadro 20 – Abreviações adotadas durante os estágios de desenvolvimento de publicações ISO	150
Quadro 21 – Detalhamento da codificação dos estágios de um documento ISO	151
Quadro 22 – Abreviações adotadas pela ISO ao longo dos diferentes estágios de elaboração dos documentos de padrões internacionais	152
Quadro 23 – Grupos de Trabalho do TC159/SC4 da ISO	154
Quadro 24 – <i>ISO 9241</i> – Partes, Grupos de Trabalho e Status Atual	154
Quadro 25a – <i>ISO 9241</i> – Síntese do conteúdo das <i>Partes 1 a 9</i> e alvos correlatos	156
Quadro 25b – <i>ISO 9241</i> – Síntese do conteúdo das <i>Partes 10 a 17</i> e alvos correlatos ...	157
Quadro 26 – <i>ISO 9241</i> – Essência das partes constituintes	158
Quadro 27 – Aspectos do contexto do usuário e da organização, das tarefas e do sistema para diferentes estilos de interação, segundo a visão da ISO	161
Quadro 28 – Síntese descritiva dos métodos apontados para a determinação da aplicabilidade/adoção de recomendações particulares do <i>ISO 9241</i>	167
Quadro 29 – Discriminação do N° de Itens do <i>IsoMetrics</i> por Princípio do <i>ISO 9241-10</i>	185
Quadro 30 – Características relativas ao usuário inclusas no <i>DePerUSI</i>	193
Quadro 31 – Aspectos gerais do ensaio	202
Quadro 32 – Aspectos específicos relativos à sondagem de campo	202
Quadro 33 – Síntese do planejamento da sondagem do usuário	203

Quadro 34 – Aspectos específicos relativos ao teste de usabilidade em ambiente controlado	204
Quadro 35 – Síntese do planejamento das tarefas de teste	205
Quadro 36 – Recursos materiais utilizados no ensaio de usabilidade	207
Quadro 37 – Tipos de Dados Coletados	213
Quadro 38 – Compilação das recomendações das Partes 14, 15 e 16 do ISO 9241 não adotadas pelo MATLAB v. 5.3.0	215
Quadro 39 – Exemplos de palavras de comando do MATLAB v. 5.3.0 que contrariam a recomendação 5.1.5 da ISO 9241-15 [ISO97b]	222
Quadro 40 – Sumário de Falhas e Parecer sobre o MATLAB v. 5.3.0 com base na <u>Inspeção de Conformidade</u> com as Partes 14, 15 e 16 do ISO 9241	238
Quadro 41 – Indicadores Quantitativos Considerados e Símbolos Correlatos	241
Quadro 42 – Aspectos gerais do ensaio	254
Quadro 43 – Variáveis selecionadas para o teste <i>F ANOVA</i> fator único	254
Quadro 44 – Síntese gráfica dos resultados do <i>procedimento Tukey-Kramer</i>	258
Quadro 45 – Síntese da classificação adotada por Queiroz e Turnell [Quei96] para problemas identificados em interfaces homem-máquina	269
Quadro 46 – Sumário de Falhas e Parecer sobre o MATLAB v. 5.3.0 com base na <u>Mensuração do Desempenho do Usuário</u>	270
Quadro 47 – Síntese dos resultados da sondagem do universo amostral com o <i>DePerUSI</i>	272
Quadro 48 – Síntese dos resultados da sondagem do universo amostral com o <i>OpUS</i>	285
Quadro 49 – Síntese da computação dos itens do <i>OpUS</i> , por seção e por categoria de usuário	288
Quadro 50 – Síntese gráfica dos resultados do <i>procedimento Tukey-Kramer</i> , em termos das seções do <i>OpUS</i> e das categorias de respondentes	294

Quadro 51 – Sumário de Problemas e Parecer sobre o <i>MATLAB</i> v. 5.3.0 com base na <u>Sondagem da Satisfação do Usuário</u>	297
Quadro 52 – Confrontação dos resultados obtidos a partir dos diferentes enfoques avaliatórios adotados	299
Quadro 53 – Lista de problemas de usabilidade identificados a partir dos três enfoques avaliatórios adotados	301

LISTA DE TABELAS

Tabela 1 – Taxas de Adoção do <i>MATLAB</i> v. 5.3.0 às Partes 14, 15 e 16 do <i>ISO 9241</i>	236
Tabela 2 – Indicadores Quantitativos Medidos	240
Tabela 3 – Valores Mínimos e Máximos dos Indicadores Quantitativos Medidos	242
Tabela 4 – Número de indivíduos associados aos mínimos nulos registrados em função das categorias pré-definidas de usuários de teste	244
Tabela 5 – Sumário da análise descritiva dos dados coletados	252
Tabela 6 – Matrizes de correlação	253
Tabela 7 – Resultados do Teste <i>F ANOVA Fator Único</i> para as variáveis selecionadas	256
Tabela 8 – Procedimento <i>Tukey-Kramer</i> para comparações múltiplas de médias	257
Tabela 9 – Resultados da aplicação do teste <i>F ANOVA fator único</i> às três seções do <i>OpUS</i>	292
Tabela 10 – Resultados da aplicação do procedimento <i>Tukey-Kramer</i> às três seções do <i>OpUS</i>	293

LISTA DE ANEXOS

Anexo A – Enfoques Avaliatórios de Interfaces com o Usuário	353
Anexo B – Membros Nacionais da ISO e da IEC	354
Anexo C – Padrões ISO para a Ergonomia da Interação Homem-Máquina	358
Anexo D – Produto Avaliado: <i>MATLAB v. 5.3.0 (The MathWorks Inc.)</i>	362
Anexo E1 – Diagrama de Seções e Sub-seções do <u>ISO 9241-14</u>	364
Anexo E2 – Diagrama de Seções e Sub-seções do <u>ISO 9241-15</u>	365
Anexo E3 – Diagrama de Seções e Sub-seções do <u>ISO 9241-16</u>	366
Anexo F1 – Lista Resultante da Inspeção de Conformidade do <i>MATLAB v. 5.3.0</i> ao <u>ISO 9241-14</u>	367
Anexo F2 – Lista Resultante da Inspeção de Conformidade do <i>MATLAB v. 5.3.0</i> ao <u>ISO 9241-15</u>	369
Anexo F3 – Lista Resultante da Inspeção de Conformidade do <i>MATLAB v. 5.3.0</i> ao <u>ISO 9241-16</u>	371
Anexo G – Versão .xls do <i>DePerUSI</i>	373
Anexo H – Versão .xls do <i>OpUS</i>	374
Anexo I – Ficha Cadastral de Participação em Ensaios de Usabilidade	375
Anexo J – Roteiros de Tarefas – Versão do Usuário	379
Anexo K – Roteiros de Tarefas – Versão do Avaliador	387
Anexo L – Fichas de Registro de Eventos	397
Anexo M – Tópicos de Estatística Indutiva	399

1

CONSIDERAÇÕES INICIAIS

"To the user of an interactive system, communication is at least as important as computation."

[Deborah Hix & Rex Hartson, *Developing User Interfaces* (1993)]

1	Considerações Iniciais	1
1.1	Introdução.....	3
1.2	Caracterização do Objeto da Pesquisa	7
1.3	Motivação e Justificativa da Pesquisa	14
1.4	Objeto da Pesquisa.....	17
1.5	Objetivos da Pesquisa.....	19
1.5.1	Geral.....	19
1.5.2	Específicos.....	19
1.6	Encaminhamento Metodológico.....	19
1.7	Organização do Documento.....	21

O presente capítulo oferece uma visão geral dos processos avaliatórios de produtos interativos usuário-computador, focalizando a atenção no objeto de estudo desta pesquisa e apresentando, ao final, a estrutura dos demais capítulos.

A Seção 1.1 (**Introdução**) ressalta, inicialmente, a importância do ato de *avaliar* no cotidiano atual e, em especial, nos processos interativos usuário-computador. Em seguida, enfatiza a relevância da interface do usuário e de sua avaliação, apontando dificuldades concernentes à sistematização dos processos destinados a sua otimização e à determinação de sua eficiência ou deficiência. Por fim, aborda as duas grandes categorias de métodos de avaliação consagrados na literatura da área.

Na seção seguinte, **Caracterização do Objeto da Pesquisa**, procura-se reforçar a significância da avaliação no contexto da concepção e desenvolvimento de produtos interativos, desde a prancheta do projetista até o uso integral das potencialidades do produto final. Destaca-se também a avaliação da *usabilidade* do produto como um processo que vem ganhando um espaço cada vez mais definido no domínio da comunicação mediada por computadores, apresentando diferentes concepções de conceituação do termo. Em seguida, enumera-se algumas razões para o crescimento da importância deste atributo em diferentes contextos de projeto e avaliação de produtos interativos e comenta-se alguns dos indicadores desse crescimento.

A Seção 1.3, **Motivação e Justificativa da Pesquisa**, apresenta razões que estimularam a realização deste trabalho de pesquisa, assim como argumentos que respaldaram a importância de sua realização.

Na Seção 1.4, **Objeto da Pesquisa**, comenta-se o tema e enuncia-se as hipóteses que fundamentaram a pesquisa, enquanto na Seção 1.5 - **Objetivos da Pesquisa** – explicita-se o objetivo geral e os objetivos específicos que nortearam a realização do trabalho.

A Seção 1.6, **Encaminhamento Metodológico**, oferece uma visão geral da metodologia adotada durante o desenvolvimento do trabalho, a qual será apresentada e discutida mais detalhadamente nos Capítulos 4, 5 e 6 deste documento.

Encerrando as considerações que delineiam a estruturação deste documento, a última seção deste capítulo apresenta a organização do texto, descrevendo sucintamente o conteúdo de cada um dos demais capítulos.

1.1 Introdução

O ato de *avaliar* parece ser inerente aos seres humanos. Desde os primórdios das civilizações, indivíduos avaliam a aparência, o comportamento e o desempenho de outros indivíduos, professores avaliam as habilidades, o progresso e a satisfação de seus alunos que, por sua vez, lhes avaliam a competência, o desempenho e o modo de condução dos cursos ministrados, compradores avaliam a qualidade das mercadorias que compram, pais avaliam os boletins dos filhos, profissionais avaliam ofertas de emprego e assim por diante.

Há uma tendência natural de se avaliar coisas que serão experimentadas ou usadas em algum contexto, seja um carro que se deseja adquirir, um apartamento que se irá alugar ou um sistema computacional que será empregado no desenvolvimento de uma atividade qualquer. Na verdade, o ato de *avaliar* parece ter sido naturalmente incorporado ao cotidiano das sociedades, tendendo a assumir novas conotações, à medida que novos contextos o exigem.

No domínio da interação homem-máquina e, mais especificamente, no sub-domínio dos processos interativos usuário-computador, onde se insere o objeto desta pesquisa, a avaliação ocorre de formas as mais diversas, podendo ser de natureza *informal* ou *formal*, *sistemática* ou *empírica*, *ad hoc* ou *rigorosa*, passível de produzir resultados subjetivos, objetivos ou combinações de ambos.

Em um contexto mais científico, avaliações consistem, em essência, de procedimentos de averiguação do se e do porque um objeto de estudo (e.g., um artefato, um manual, um procedimento) preenche ou não os requisitos dele exigidos ou a ele atribuídos [Marc94, Quei96, Quei97a, Quei97b, Quei98a]. Os métodos empregados dependem de quando, porque e por quem a avaliação é conduzida, assim como da gama de limitações contextuais e dos recursos disponíveis no contexto considerado [Pree90, Hix93, Quei96, Quei98a, Harv98, Mayh99].

A avaliação constitui uma etapa essencial e integral do projeto e desenvolvimento de sistemas de informação computadorizados [Pree90, Quei94, Quei96, Dix98]. O desempenho de sistemas computacionais tem sido avaliado desde a época do modo de processamento por "lotes" (*batch processing*), a partir da seleção e execução de programas, com propósitos de mensuração de parâmetros de interesse (e.g., número de ciclos de CPU, número de acessos ao HD) para análise posterior [Kant92].

O advento dos sistemas computacionais com compartilhamento de tempo (*time-sharing*) passou a exigir iniciativas de avaliação mais complexas. Assim, foram desenvolvidos roteiros contextuais (*contextual scenarios*) interativos que eram empregados, sob condições controladas, para a ativação de sistemas computacionais, a partir de um conjunto de terminais, enquanto se mensurava o desempenho do sistema. Os esforços despendidos, no entanto, enfatizavam sobretudo o desempenho "interno" do sistema computacional, i.e. apenas a avaliação da componente sistema computacional, independentemente de *como* os resultados obtidos exerciam impactos sobre o usuário. Tais estratégias limitavam seu alcance para o processo interativo usuário-computador, já que eram relevantes apenas para projetistas de *hardware* e *software* computacionais, porém pouco expressivas para o usuário envolvido no processo interativo [Treu94].

Vale a pena salientar que, no tocante aos sistemas interativos usuário-computador, distingue-se fundamentalmente quatro componentes interagentes - o **suporte computacional físico** (*hardware*), o **suporte computacional lógico** (*software*), o **usuário** e a **interface usuário-computador**¹ (também referida como **interface com o usuário** ou simplesmente **interface**), que proporciona o elo de comunicação entre as demais e, sem dúvida, representa uma das componentes mais importantes (se não a mais importante) do processo de comunicação mediada por sistemas computacionais [Bø82, Neel90, Mayh92, Hix93, Quei96, Quei97a, Mand97, Dix98].

Sendo o meio através do qual o usuário se comunica com um sistema computacional, a interface usuário-computador exerce uma influência significativa no modo *como* o usuário visualiza e entende a funcionalidade de um sistema, o que vincula sua concepção a aspectos do sistema com os quais o usuário interage física e cognitivamente. O ISC [ISC97] considera a *interface* como a "cara" do sistema com o qual um indivíduo (o usuário) interage, acrescentando que é necessária a existência de sistemas mais fáceis de usar por quaisquer indivíduos, caso a meta a atingir seja a comunicação eficaz entre homens e sistemas computacionais, e que uma comparação entre as facilidades atuais de comunicação homem-máquina com as formas de comunicação homem-homem mostra que as interfaces homem-máquina atuais são extremamente limitadas.

Segundo Hackos e Redish [Hack98], a *interface* é o que os usuários vêem ao usar o produto e o meio através do qual se comunicam com o produto ao usá-lo, corroborando a visão de Hix e Hartson [Hix93] de que a *interface* é, para os usuários, "o sistema".

No caso de aplicações de *software*, Mayhew [Mayh99] acrescenta que a *interface* pode ser definida como as linguagens através das quais o usuário e o produto se comunicam. Processos de comunicação desta natureza estão usualmente associados tanto com o modo como a apresentação e o retorno da informação são projetados (linguagem da aplicação para o usuário) quanto com o modo como os usuários indicam para a aplicação o que desejam fazer em seguida a partir de interações com dispositivos de visualização - em geral, terminais de vídeo - através de dispositivos de entrada, tais como teclado e *mouse* (linguagem do usuário para a aplicação).

De um modo mais abrangente, que transcende as aplicações de *software*, o ISC² [ISC97] define a *interface* de um sistema como a componente que engloba os vários meios através dos quais os usuários se comunicam e interagem com o sistema, a natureza do diálogo e os meios através dos quais o sistema comunica suas respostas aos usuários, incluindo o *hardware* e o *software* de entrada e saída das mais diferentes naturezas envolvidos no processo interativo, além de características de aplicação, tais como o gerenciamento e a apresentação da informação, aspectos de processos interativos cooperativos suportados por computadores (CSCW) e do compartilhamento de tarefas e responsabilidades entre usuários e sistemas (e.g., por meio de agentes autônomos).

Além disso, a visão do ISC também contempla as facilidades técnicas (e.g., comunicações, computação) que suportam a interação entre usuários e sistemas também fazem parte da interface, tendo em vista que o desempenho de tais facilidades exerce uma influência direta nas experiências dos indivíduos que utilizam sistemas interativos com propósitos os mais

¹ Neste documento os três termos serão empregados indistintamente.

² Infrastructure Steering Committee do National Research Council dos EUA.

diversificados. Além de ser uma visão mais sinóptica da questão, tal definição também leva em conta os contextos em que se inserem os processos interativos. Por outro lado, expressa a preocupação com os "lapsos" advindos da especificidade das definições dos termos *interface* e *aplicação*, onde se tende, não raramente, a omitir ou distorcer aspectos importantes da integração *interface-aplicação*, criando uma lacuna indesejável na concepção do todo.

A atenção crescente dada às interfaces usuário-sistema, principalmente a partir de meados da década de 80, assim como o cuidado em considerar aspectos os mais abrangentes em sua definição, só vem reforçar o aumento da importância que o processo de comunicação entre indivíduos e máquinas (e, mais particularmente, entre usuários e sistemas computacionais) vem adquirindo no que concerne à entrada, manipulação e saída de informações.

Além disso, constata-se também que indivíduos atuantes nas mais diferentes áreas do conhecimento humano vêm se tornando, em número cada vez mais elevado, usuários (e, por conseguinte, dependentes [Pree95]) da tecnologia da informação, a cada dia mais indispensável e integrante do cotidiano atual. Entretanto, o uso cotidiano vem contribuindo significativamente para enfatizar que tal condição de dependência gera, em algum estágio do uso das tecnologias computacionais, sentimentos de frustração ou dúvida em usuários dos mais diferentes perfis, engajados em tarefas de naturezas as mais variadas [ISC97].

Este cenário é reforçado pela evolução dos mercados de sistemas computacionais em ritmos consideravelmente superiores àqueles dos demais setores industriais. Tal evolução sofre fortes influências de diferentes fatores, dentre os quais podem ser mencionados a necessidade de redução de custos e a demanda por aplicativos cada vez mais complexos, poderosos e com melhor desempenho [Bias94, Rodr98, Shne00].

Não restam dúvidas de que os sistemas interativos fazem parte do cotidiano de praticamente todos os ramos de atividades do mundo contemporâneo. Sua importância sócio-econômica tem crescido em ritmo cada vez mais vertiginoso, assim como sua complexidade, diversidade e aplicabilidade, tornando as "soluções" atuais obsoletas em futuros cada vez mais próximos, o que exige dos desenvolvedores novas soluções em prazos cada vez mais curtos. Tais fatores, por sua vez, vêm redirecionando as atividades de projeto, implementação e avaliação de produtos de *hardware* e *software* destinados ao uso cotidiano por um sem número de usuários de sistemas computacionais, dotados de diferentes graus de formação, experiências, habilidades e expectativas.

Eis porque o somatório de todos estes aspectos tem contribuído, em maior ou menor grau, principalmente nas últimas duas décadas, para que a comunidade especialista em processos interativos focalize seus esforços em um sem número de iniciativas centradas na compreensão do papel do usuário e dos modos mais adequados de interação deste com sistemas computacionais. No âmbito desta realidade, diversos estudos vêm procurando enfatizar cada vez mais a relevância do processo de comunicação entre aplicações e usuários, assim como a importância dos procedimentos avaliatórios durante o ciclo de vida de produtos interativos e, em especial, o papel da *usabilidade* como atributo de projeto e avaliação (e.g., Spencer [Spen85], Whiteside [Whit88], Karat *et al.* [Kara92], Shneiderman [Shne93], Nielsen [Niel93c], Treu [Treu94], Rubín [Rubi94], Henry [Henr98], Hackos e Redish [Hack98], Dix *et al.* [Dix98], Mayhew [Mayh99], [Shne00]).

Deste modo, várias técnicas³ de desenvolvimento e avaliação de processos interativos, fundamentadas em *modelos mentais do usuário*, na *estruturação cognitiva de novas informações pelo cérebro humano*, em *contextos de trabalho do usuário* e na *engenharia da usabilidade*⁴, vêm sendo concebidas por equipes de pesquisa e adotadas por corporações de desenvolvimento e avaliação de produtos de *software* de naturezas as mais diversas.

Entretanto, apesar de todas essas iniciativas, muitos dos esforços envidados pelos grupos de desenvolvimento de sistemas interativos ainda resultam em modos relativamente pouco sistemáticos de tornar as interfaces usuário-computador melhores [Thim94, Thim99, Thim00]. Adicionalmente, apesar dos problemas de usabilidade registrados em interfaces desenvolvidas para o consumidor serem bem conhecidos, os desenvolvedores de produtos ainda continuam a disponibilizar equipamentos e aplicações de *software* se se preocuparem com o quão eficientes ou deficientes o são [Thim94]. Thimbleby [Thim2000] assume que, sobretudo no caso de interfaces de eletrodomésticos (e.g., fax, videocassetes), isto ainda ocorre por conta da miscelânea de características e de botões que confundem os usuários e os impede de avaliar convenientemente as interfaces dos eletrodomésticos no curto espaço de tempo em que se encontram na loja efetuando a escolha.

Tal afirmação só reforça a necessidade de uma maior preocupação com a questão da avaliação. Diversos métodos de avaliação de processos interativos têm sido propostos e testados nas duas últimas décadas. Grosso modo, tais métodos podem ser enquadrados em duas grandes categorias, as quais foram rotuladas e descritas em maiores detalhes por Treu [Treu94] e serão revisadas e atualizadas no próximo capítulo deste trabalho.

Segundo a classificação de Treu [Treu94], também adotada neste trabalho, as duas grandes famílias de métodos avaliatórios podem ser sintetizadas em:

(i) *métodos de avaliação da usabilidade*, desenvolvidos desde meados dos anos 80 e centrados na análise de aspectos relacionados com a usabilidade de interfaces usuário-computador, são resultantes dos esforços de inúmeros pesquisadores, e.g., Spence [Spen85], Whiteside [Whit88], Nielsen e Molich [Niel90], Karat [Kara92], Shneiderman [Shne93], Nielsen [Niel93a, Niel94a], Treu [Treu94], Dumas e Redish [Duma94], Hartson *et al.* [Hart96], Doubleday *et al.* [Doub97], Zhang *et al.* [Zhan98a, Zhan98b] e Costabile e Matera [Cost99]; e

(ii) *métodos empíricos e outros métodos de avaliação*, resultantes de estudos experimentais que envolvem *comparações* entre técnicas e/ou estilos de interação⁵, os quais também vêm sendo desenvolvidos e/ou validados desde meados da década de 80 por um

³ Treu [Treu94] distingue os termos *método* e *técnica*, alegando que uma *técnica* é empregada como um suporte de uma ou mais etapas num procedimento mais geral denominado método. Logo em seguida, no entanto, o autor ressalta que ambos são muitas vezes empregados indistintamente. É o que se percebe confrontando os textos de Jeffries *et al.* [Jeff91] e Miller e Jeffries [Mil92] - onde o termo empregado é *técnica* - com o texto de Treu [Treu94] (que inclui os estudos de Jeffries *et al.* [Jeff91]), no qual o termo *método* é constantemente utilizado para descrever a mesma categorização. Neste documento, os termos em questão serão empregados indistintamente.

⁴ O termo *Engenharia da Usabilidade*, questionável em português por alguns puristas, assim como o é o termo *Engenharia de Software*, será discutido mais detalhadamente em uma das seções do próximo capítulo.

⁵ Vide Seção 2.3 (*Métodos Empíricos e Outros Métodos de Avaliação*), onde se aborda mais detalhadamente os métodos agrupados sob este rótulo.

segmento representativo da comunidade pesquisadora da área de interação usuário-computador, e.g., Jeffries e Rosenberg [Jeff87], Callahan [Call88], Card *et al.* [Card90], Morgan *et al.* [Morg91], Gray *et al.* [Gray92], Eberts e Bittianda [Eber93], Benbasat e Todd [Benb93], Trumbly *et al.* [Trum93], Brewster *et al.* [Brew94], Zanden e Venckus [Zand96], Watson *et al.* [Wats97], Zhang e Ren [Zhan97], Scholtz [Scho99], Koskinen [Kosk99].

1.2 Caracterização do Objeto da Pesquisa

A *avaliação* não constitui apenas um dos estágios integrantes do processo de projeto, nem tampouco uma série de atividades inseridas fortuitamente no estágio final do processo de implementação, caso as restrições temporais ainda o permitam.

A *avaliação* é parte integrante do processo de projeto, estendendo-se ao longo de todo o ciclo de vida de desenvolvimento da interação usuário-computador [Hix93, ISC97, Mayh99]. O projeto e a *avaliação* de produtos estão estreitamente vinculados em todos os níveis, desde o teste informal de idéias de projeto ao teste formal de protótipos e sistemas implementados [Henr98]. A questão central associada ao processo avaliatório é *como* melhor avaliar a qualidade de uso de um sistema por seus usuários reais ou hipotéticos, executando tarefas em seu ambiente de trabalho. Uma vez que haja um protótipo disponível, a *avaliação* poderá ser centrada no usuário [MacI96, Liu97, Dix98, Mayh99]. A questão central associada ao processo avaliatório é *como* melhor avaliar a qualidade de uso de um sistema por seus usuários reais ou hipotéticos, executando tarefas em seu ambiente de trabalho.

Os processos avaliatórios visam testes da funcionalidade e usabilidade de projetos, assim como a identificação e a correção de problemas de quaisquer naturezas, sendo conduzidos usualmente em laboratórios ou ambientes de trabalho do usuário e envolvendo, em maior ou menor grau, sua participação ativa.

Idealmente, a *avaliação* deveria ocorrer ao longo do ciclo de vida do produto, seus resultados servindo de respaldo para otimizações do produto considerado. No entanto, nem sempre é possível a condução de ensaios avaliatórios de modo contínuo ao longo de todo o ciclo de vida do produto. Porém, é possível (e sempre pertinente) empregar técnicas analíticas e informais durante as fases de projeto e implementação do produto.

Neste sentido, as técnicas de *avaliação*, de modelagem e de prototipagem possuem vínculos estreitos entre si. Tais técnicas auxiliam o desenvolvedor a assegurar a *avaliação* contínua do projeto, antecipando a solução de problemas ao dispêndio de esforços e recursos com a implementação, muitas vezes desnecessários e significativos. Na realidade, é muito mais fácil alterar um projeto nas etapas iniciais de desenvolvimento do que em seus estágios finais.

É possível estabelecer uma distinção nítida entre o processo de *avaliação* do projeto de um sistema interativo e os processos de *avaliação* de uma implementação desse projeto, quer no estágio de protótipo ou de produto final. No primeiro caso, a *avaliação* ocorre tipicamente segundo o ponto de vista do projetista, sem o envolvimento direto da comunidade usuária do produto. Por outro lado, quando a *avaliação* é conduzida durante a etapa de implementação do projeto, o foco das análises é direcionado para o uso real do sistema [Dix98].

Um produto pode ser avaliado mesmo antes de haver sido tomada qualquer decisão de implementação, com propósito de minimizar gastos com erros de projeto no estágio preliminar de desenvolvimento. A maioria das técnicas para avaliação empregadas neste estágio é analítica, o que, via de regra, implica o envolvimento de especialistas no processo avaliatório do projeto, fundamentados em princípios cognitivos e de usabilidade.

No estágio preliminar de desenvolvimento, resultados experimentais prévios e procedimentos de modelagem de usuários e tarefas também podem proporcionar valiosas contribuições ao desenvolvimento do projeto. Uma vez desenvolvido um protótipo ou uma versão completa do produto, técnicas experimentais e observatórias poderão ser empregadas para a aquisição de resultados quantitativos e/ou qualitativos.

O processo de seleção do método de avaliação de um produto está intimamente vinculado ao *que se requer* como resultado da avaliação. Daí, a existência de uma grande variedade de métodos de avaliação, cada um aplicável a diferentes estágios e contextos de projeto. Alguns se afiguram mais subjetivos do que outros, possibilitando a coleta de informações qualitativas, ao invés de medidas quantitativas. Alguns fornecem informação imediata, ao passo que outros favorecem processos posteriores de realimentação de informações. Finalmente, alguns exigem mais recursos temporais, físicos, materiais e humanos do que outros, conforme será discutido no Capítulo 3.

A experiência com avaliação da usabilidade tem mostrado que tanto no projeto de objetos de interação quanto na forma de sua incorporação nos diálogos que auxiliam as tarefas de um contexto de trabalho, até mesmo alterações insignificantes na apresentação e realimentação da informação podem exercer uma influência significativa na qualidade de interação [MacI96]. Estratégias de avaliação de sistemas computacionais que não consideram como os seus resultados influenciarão as ações e o desempenho futuros do usuário apresentam uma importância limitada no contexto da interação usuário-computador. Embora estas estratégias sejam de interesse para projetistas e desenvolvedores de *hardware e software*, não são em si relevantes do ponto de vista do usuário, a menos que o objetivo seja verificar efeitos do funcionamento interno da máquina no processo comunicação externo [Treu94, Quei96, Henr98].

A *avaliação* de sistemas computacionais e, mais especificamente, de interfaces usuário-computador possibilita a injeção de objetividade no processo de análise de um projeto e se aplica ao sistema como um todo, incluindo não só suas componentes tecnológicas, mas também a componente humana [Hix93, Sher97, Stan97a, Henr98], tão importante no contexto de qualquer processo interativo homem-máquina.

Como o desempenho de qualquer sistema computacional que envolve a interação com seres humanos não depende somente do desempenho da máquina, o processo de avaliação do sistema implica a análise da "parceria" usuário-computador [Card90], i.e., o estudo analítico de usuários interagindo com os recursos oferecidos [Down91, Sioc91, Cox93, Park93, Treu94, Liu97, Hack98]. Só assim serão refletidos nos resultados características psicológicas dos usuários e aspectos relativos à realização da tarefa utilizando o sistema computacional.

No entanto, a despeito da atenção a cada dia crescente com que a concepção e implementação de interfaces usuário-computador vêm sendo encarados por diversas organizações

de desenvolvimento de *software*, o projeto e a *avaliação* satisfatória de novas interfaces usuário-computador ainda apresentam "gargalos" difíceis de solucionar [Gong94, Thim99, Shne00], sobretudo quando ainda se constata:

(i) a dificuldade de quantificação e justificativa das relações custos-benefícios da implantação de programas de usabilidade e dos processos avaliatórios da usabilidade dos produtos desenvolvidos [Ehrl94a, Lyne95, Rous97, Kade98];

(ii) a heterogeneidade do público alvo das recomendações de usabilidade, em geral uma miscelânea de engenheiros de *software*, administradores, gerentes de projetos, escritores técnicos e profissionais afins das mais diferentes competências, carentes de informações com graus de detalhamento os mais diversos [Naya95];

(iii) a visão distorcida das equipes de desenvolvimento que considera os programas de usabilidade "extras opcionais" que só oneram e prolongam os ciclos de vida dos projetos [Hix93, Kade98, Mayh99];

(iv) a necessidade de reversão da percepção e postura ambivalentes dos gerentes de projeto no que diz respeito à relevância da implantação de programas de usabilidade em seus setores [Mayh94, Kade98, Mayh99];

(v) a relutância (e por que não dizer a suspeita, a oposição e, muitas vezes, a animosidade) das equipes de desenvolvimento face aos agentes da coleta de dados e às críticas e recomendações das equipes de avaliação [Naya95, Kade98];

(vi) o acesso limitado, por inúmeras circunstâncias (e.g., carência de motivação da parte dos desenvolvedores, dificuldade de "mapeamento" de comunidades usuárias em potencial, resistência dos gerentes de grupos de usuários, distâncias geográficas, barreiras sócio-culturais) às comunidades usuárias [Mayh94, Kade98];

(vii) a inércia das organizações no que concerne à oferta de produtos usáveis, em parte "estimulada" pelo posicionamento ainda pouco exigente dos usuários quanto à qualidade e usabilidade dos produtos que adquirem [Kade98].

Alie-se a estas constatações a relevância do entendimento apropriado dos desenvolvedores sobre quais as técnicas que se adequam a diferentes tarefas e estilos de interação considerados, assim como a diferentes contextos de trabalho, face às inúmeras abordagens avaliatórias desenvolvidas ao longo das últimas décadas pela comunidade pesquisadora de processos interativos. Tal entendimento é imprescindível na tomada de decisão de *qual* abordagem se enquadra melhor no contexto de trabalho considerado, visto que, embora convergentes em nível de solução de problemas, diferentes abordagens avaliatórias apresentam diferentes potencialidades e limitações.

Por fim, acrescente-se o crescimento em progressão geométrica do universo de usuários potenciais de sistemas computacionais e o redobramento conseqüente da atenção das equipes de desenvolvimento e avaliação à concepção e ao desenvolvimento de produtos cuja aceitabilidade social e prática possa implicar uma inquestionável usabilidade de suas interfaces usuário-computador [Niel93c, Mayh99, Thim99]. Afinal de contas, a elevação dos níveis de confiança e satisfação dos usuários de produtos de *software* dependerá da atenção dada à maturidade dos

processos de desenvolvimento de tais produtos, assim como do ponto de vista dos processos internos de desenvolvimento [Curt93] e avaliação.

Inserido neste cenário, o estudo da usabilidade vem ocupando, de forma cada vez mais bem delineada, nos domínios do projeto centrado no usuário [Hix93, Mand97, Henr98, Kein99] e da engenharia da usabilidade [Whit88, Niel92a, Niel93a, Niel93b, Niel93c, Duma94, Rubi94, Mayh99, Thim99], um espaço próprio no contexto das atividades de desenvolvimento de *software*, constituído por um conjunto de métodos e abordagens de projeto e avaliação. Sua importância se consolida a cada dia no domínio da concepção de bens de consumo de naturezas as mais diversas, em especial produtos computacionais [Wiki94, Jord96a, Jord96b], fazendo com que a disponibilidade de um produto no mercado consumidor torne-se uma condição limitante absoluta de sua usabilidade, já que, do ponto de vista da comunidade usuária, este aspecto implica, em primeiro lugar, a oferta de aplicações que possam atender às suas necessidades e, em segundo lugar, que sejam convenientemente usáveis em atendimento às suas necessidades [Albe94, Curt94, Jord96b, Hack98, Mayh99].

No contexto de estudo da engenharia da usabilidade, o termo *software amigável*, bastante difundido nos anos 80, apesar de impreciso e subjetivo, tem sido, pouco a pouco, substituído pela comunidade usuária de produtos de *software* pelo termo *software usável*, com uma conotação que ressalta, em geral, uma qualidade positiva do produto – a facilidade de comunicação com o usuário [Medy93]. Apesar da importância que os estudos da usabilidade vêm ganhando nas duas últimas décadas, o termo *software usável* e, por extensão, o termo *usabilidade* ainda não possuem definições precisas. As definições existentes refletem visões distintas, porém não mutuamente exclusivas, da natureza da usabilidade [Medy93, MacI96], sobretudo face à multidimensionalidade do termo [Kein99]. Segundo Keinonen [Kein99], os processos de desenvolvimento que visam assegurar a usabilidade de produtos devem incluir fases e objetivos que integrem o caráter multifacetado do termo.

Atualmente, vem se tornando particularmente popular e consistente a visão contextual, através da qual a usabilidade de um produto é estabelecida em função de universos amostrais de usuários específicos sob condições de estudo, das tarefas que estes executam e do ambiente no qual atuam [Davi93, Beva94, MacI96, ISC97, ISO98, Mayh99, Kein99].

Uma das razões que respaldam este ponto de vista é a falta de sentido inerente à tentativa de desenvolvimento e avaliação de um sistema sem a consideração do contexto em que este será empregado. Afinal de contas, a usabilidade de um produto não é unidimensional, compreendendo diversas componentes, às vezes conflitantes [Niel93b, Kein99], como será discutido nos parágrafos seguintes e, mais detalhadamente, no próximo capítulo.

Conforme o IEEE⁶ [IEEE90a], usabilidade é definida como a facilidade com que um usuário pode aprender a operar, preparar as entradas e interpretar as saídas de um sistema ou componente de sistema.

Na concepção de Shackel [Shac91], a usabilidade de um sistema ou equipamento é a característica que o capacita, em condições funcionais humanas, a ser usado fácil e efetivamente,

⁶ The Institute of Electrical and Electronics Engineers, Inc. (Instituto de Engenheiros Eletricistas e Eletrônicos).

a partir de um determinado nível de treinamento e suporte, por uma determinada gama de usuários na execução de um determinado conjunto de tarefas, no contexto de uma gama pré-definida de cenários ambientais. Em suma, a usabilidade de um sistema ou equipamento é a capacidade de permitir ser usado fácil e efetivamente por humanos.

Por outro lado, o ETSI⁷ [ETSI93] considera a usabilidade como um conceito puramente ergonômico, independente de custos de provisão do sistema que, juntamente com a razão entre os benefícios usufruídos pelo usuário e os custos financeiros despendidos, compõem o conceito de utilidade. Tal concepção conduz à constatação de que um sistema usável altamente ergonômico possa apresentar baixa utilidade para um usuário específico que considere os custos envolvidos muito elevados com relação a sua necessidade de uso.

Nielsen [Niel93b, Niel94b] considera a usabilidade como um conceito geral composto por parâmetros que podem ser mensurados. O autor afirma que a usabilidade é dos aspectos que influenciam a aceitação de um produto e sugere explicitamente que a usabilidade (*usability*), integrada à utilidade⁸ (*utility*), compõe a aplicabilidade (*usefulness*) do produto, ao afirmar que enquanto a utilidade diz respeito a se a funcionalidade do sistema satisfaz às necessidades para as quais ele foi desenvolvido, a usabilidade relaciona-se a quão bem os usuários podem utilizar aquela funcionalidade [Niel93b].

Este ponto de vista também é consonante tanto com a concepção de usabilidade de Eason [Easo84], quando este afirma que a usabilidade de um sistema pode limitar o nível de percepção da aplicabilidade potencial de um sistema computacional por seus usuários, quanto com a perspectiva de usabilidade de Grudin [Grud92], ao afirmar que as diferenças entre a usabilidade e a utilidade de um sistema se refletem nos processos de projeto, acrescentando que a utilidade é definida no início pelos gerentes de produtos, enquanto a usabilidade só é otimizada subsequente pelos projetistas. O próximo capítulo discute mais detalhadamente o modelo de usabilidade de Nielsen [Niel93b].

Preece [Pree94] relaciona a usabilidade à facilidade de aprendizado e de uso de um produto, sua segurança, efetividade e eficiência, assim como à opinião global dos usuários sobre tal produto. Dumas [Duma94], por outro lado, associa a usabilidade à rapidez e à facilidade com que os usuários de um produto executam satisfatoriamente suas tarefas.

Conforme o projeto RACE⁹ [RACE95], a usabilidade é definida a partir da perspectiva do projetista, ou seja, do que os projetistas necessitam a fim de garantir o desenvolvimento de sistemas e serviços usáveis. O projeto desenvolveu um método para a construção de uma especificação estruturada, centrada na tarefa do usuário, com base nas noções de *tarefas do usuário* (*user tasks* - o que o usuário deseja executar) e *tarefas de habilitação* (*enabling tasks* - o que o sistema requer que o usuário execute). Um conjunto de *princípios genéricos de projeto*

⁷ *European Telecommunications Standards Institute* (Instituto Europeu de Padrões para Telecomunicações).

⁸ *Usefulness* e *utility* são neologismos neste contexto, suas definições são pouco distintas, mesmo em bons dicionários da língua inglesa, e.g. *Oxford Advanced Learner's Dictionary* [Horn90]. Daí, tradução como *aplicabilidade* e *utilidade*, palavras existentes no *Novo Dicionário da Língua Portuguesa* [Holl86], que se enquadram no contexto em questão.

⁹ *Research and Development in Advanced Communications Technologies in Europe* (Pesquisa e Desenvolvimento de Tecnologias Avançadas em Telecomunicações na Europa).

(*generic design principles*) e tabelas de especificação (*specification tables*) suportam a conclusão de uma *análise de estados de habilitação* (*enabling states analysis - ESA*). O projeto também desenvolveu um protótipo computadorizado (*Design Support Tool - DSP*) para gerenciar a aplicação dos princípios e a construção e cruzamento das tabelas de especificação.

Segundo a ISO¹⁰ [ISO97a, ISO97b, ISO98, ISO99a] e o Projeto MUSIC [Beva94], usabilidade é definida como a efetividade¹¹, a eficiência¹² e a satisfação¹³ com que determinadas categorias de usuários podem alcançar metas pré-definidas em ambientes específicos. A Parte 11 do padrão internacional ISO 9241 também enfatiza que a usabilidade depende do contexto de uso e que o grau de usabilidade atingido dependerá das circunstâncias específicas nas quais um produto é usado. O *contexto de uso* consiste dos usuários, tarefas, equipamentos (*hardware, software* e materiais) e ambientes físicos e sociais (que também poderão influenciar a usabilidade de um produto em um sistema de trabalho¹⁴). Deste modo, medidas de desempenho e satisfação do usuário avaliam o sistema global de trabalho. Quando o foco de interesse é um produto, tais medidas fornecem informação sobre a usabilidade daquele produto no contexto particular de uso estabelecido pelo restante do sistema de trabalho.

É importante observar que enquanto esta definição fornece um modo prático de mensuração da usabilidade, também considera as conseqüências de outras características de qualidade de *software*, e.g. funcionalidade, confiabilidade, eficiência do sistema computacional. Alterações nestas características ou em outros componentes do sistema de trabalho, tais como o grau de treinamento do usuário ou o melhoramento da iluminação, também podem exercer um impacto no desempenho e na satisfação do usuário [Beva97].

Ao longo deste documento, o termo *usabilidade* é encarado em diferentes momentos sob três pontos de vista, a saber: (i) do desenvolvedor, focalizando quer o projeto ou o produto final; (ii) do usuário, focalizando necessidades, metas a atingir e ações necessárias para atingi-las; e (iii) do avaliador, focalizando métodos e abordagens de análise e mensuração.

Entretanto, constatar-se-á que, ao longo deste documento, o termo *usabilidade* será empregado fundamentalmente com duas conotações. A primeira delas diz respeito à facilidade de aprendizado dos recursos oferecidos por um sistema, à flexibilidade e à confiabilidade do processo de comunicação usuário-sistema, tanto no tocante à estruturação das entradas (comunicação do usuário com o sistema) quanto ao fornecimento de saídas pelos diferentes componentes do sistema (comunicação do sistema com o usuário). A segunda conotação é a de atributo inerente a um sistema, variável conforme o contexto e as circunstâncias de uso do sistema por indivíduos e mensurável através dos graus de desempenho e satisfação apresentado por esses indivíduos ao realizarem uma seqüência de procedimentos com o propósito de atingir uma série de metas pré-fixadas.

¹⁰ *International Standards Organization* (Organização para a Elaboração de Padrões Internacionais)

¹¹ Precisão e completude com que os usuários atingem metas especificadas.

¹² Recursos despendidos com relação à precisão e completude com que os usuários atingem metas.

¹³ Ausência de desconforto e exibição de atitudes positivas no tocante ao uso do produto.

¹⁴ *Work system*, nos textos do padrão internacional ISO 9241.

À primeira vista, poderá parecer que tais conotações privilegiam a perspectiva do usuário, o que em parte é verdadeiro, pois se acredita que o usuário é o centro do processo de desenvolvimento de sistemas interativos. Todavia, uma leitura mais cuidadosa deixará entrever a conexão desta perspectiva (aparentemente "privilegiada") com as demais perspectivas anteriormente mencionadas.

Vale a pena salientar que a concepção multifacetada do termo *usabilidade* é consonante com a visão de diversos autores, dentre os quais Shneiderman [Shne93], Mandel [Mand97], Kaderbhai [Kade98] e Keinonen [Kein99]. A primeira conotação empregada neste documento apresenta interseções, em maior ou menor grau, com as definições de usabilidade encontradas, por exemplo, em IEEE [IEEE90a], Hix e Hartson [Hix93], Preece [Pree94] e Dix *et al.* [Dix98], enquanto a segunda conotação é consonante com a definição contida na *Parte 11* do padrão internacional ISO 9241 [ISO98].

De tudo o que foi anteriormente exposto, apreende-se que a avaliação da usabilidade pode assumir significados bastante distintos para diferentes profissionais e contextos [Mac96]. Procedimentos de avaliação da usabilidade de sistemas de informação e de tecnologias interativas podem focar questões específicas ou abrangentes. Podem envolver estudos de conformação do produto com diretrizes de projeto e padrões, refletir opiniões de especialistas sobre o produto, concentrar-se na satisfação subjetiva do usuário no tocante ao produto ou em seu desempenho ao usá-lo. Seu escopo pode abranger desde o estudo da interação de usuários individuais com aspectos detalhados de uma interface à análise do uso de um sistema de suporte ao trabalho cooperativo desenvolvido em nível organizacional.

O estilo, o cronograma, o dimensionamento e o planejamento de qualquer estudo avaliatório são aspectos intimamente vinculados aos objetivos almejados, assim como à seleção dos indivíduos envolvidos. Na prática, a natureza dos processos avaliatórios, em alguns casos, parece ser ditada (i) pelos recursos materiais, físicos e humanos disponíveis; (ii) pelas preferências da equipe de avaliação por estratégias particulares de ação; (iii) ou pela falta de conhecimento dos gerentes de processos no que diz respeito à otimização de soluções de compromisso dos aspectos envolvidos para atingir seus objetivos.

Grosso modo, os processos avaliatórios visam assegurar o alcance das metas de usabilidade (métodos para a *usabilidade*) ou de outros critérios (métodos *empíricos*) pré-determinados, bem como identificar as alterações necessárias à otimização do produto considerado sob a ótica de projeto adotada [Treu94]. Tais propósitos irão servir de ponto de partida para atividades de reprojeto da interface, bem como de foco para a concepção de novas abordagens avaliatórias.

A última década tem testemunhado notáveis alterações nas perspectivas de projeto, avaliação e uso de sistemas interativos. Na verdade, o campo da interação usuário-computador sofreu um processo de reorientação, passando de uma componente de pouca importância da engenharia de *software* para um foco de atenção de pesquisadores atuantes em inúmeras áreas de concentração, inclusive a psicologia e a ciência social. A literatura da área atesta que a prática da avaliação da usabilidade de processos interativos homem-máquina tem se expandido consideravelmente, especialmente nos últimos oito anos.

Graças ao processo de informatização de ambientes de pesquisa pura e aplicada nas mais diversas áreas do conhecimento humano, a *usabilidade* vem se tornando, cada vez mais efetivamente, um eixo-condutor das atividades do mercado de trabalho [Curt94], assim como do mercado de consumo de produtos de *software*, no que concerne a tomadas de decisão para a aquisição de ferramentas computacionais de uso cotidiano.

Dentre as diferentes razões que podem ser especuladas para a explicação deste fato sobressaem: (i) a necessidade premente de otimização de processos e produtos interativos, em virtude do aumento da importância do aspecto *usabilidade* como critério de aceitação do produto final pela comunidade usuária; (ii) a relativa facilidade com que a etapa de avaliação pode ser inserida no ciclo de vida de um processo/produto, sem alterá-lo substancialmente em nível de planejamento e gerenciamento; (iii) os benefícios advindos da implantação de um programa de usabilidade, face aos custos envolvidos; e (iv) a diversidade de estratégias avaliatórias concebidas, aplicáveis aos vários estágios de ciclos de vida de desenvolvimento de diferentes produtos.

Um indicador significativo do crescimento da importância do papel da usabilidade em atividades de desenvolvimento de produtos de *software* é o aumento dos orçamentos destinados à prática da engenharia da usabilidade [Levi96a] e, por extensão, do número de ambientes dedicados à avaliação da usabilidade e, de um modo mais abrangente, ao desenvolvimento de processos avaliatórios em geral [Niel94b].

Em 1971, Shackel [Shac71] estimou que um percentual razoável para orçamentos de usabilidade de sistemas não militares era em torno de 3%. Dezoito anos mais tarde, Wasserman [Wass89, in Niel93a] identificou, a partir de um estudo de várias corporações, que as companhias de maior expoente destinavam aproximadamente de 4 a 6% de seus quadros de pesquisa e desenvolvimento ao projeto de interfaces usuário-computador e à prática da usabilidade. Em 1993, Nielsen [Niel93c] realizou um levantamento com base em 31 projetos e determinou que o percentual mediano de seus orçamentos atribuídos à engenharia da usabilidade era 6%, o que representa uma estimativa de crescimento de 100% em cerca de 20 anos!

1.3 Motivação e Justificativa da Pesquisa

Nos dias atuais, os produtos utilizados em casa e nos ambientes de trabalho vêm se tornando cada vez mais complexos em termos de características e funcionalidades, tornando imprescindível a focalização dos esforços dos desenvolvedores sobre as necessidades e limitações daqueles que os utilizarão em suas atividades cotidianas. Caso contrário, a relação entre os benefícios oferecidos e os custos despendidos na obtenção de tais benefícios poderá tornar-se pouco compensadora para a comunidade usuária.

A usabilidade de produtos tem se tornado, especialmente nos últimos dez anos, um fenômeno empresarial e internacional - um aspecto importante no que diz respeito ao *marketing* e às vendas de produtos, tendo em vista o interesse cada vez maior das grandes, médias e pequenas corporações de inúmeros países na difusão de seus produtos pelo mundo inteiro. Eis porque a importância do desenvolvimento de produtos usáveis vem recebendo uma atenção crescente do setor industrial que, pouco a pouco, tem passado a encarar a usabilidade de seus produtos como parte da responsabilidade da indústria com seus consumidores e, por extensão, como um fator importante no sucesso das vendas [Jord96a, Jord96b, Kade98].

A questão da internacionalização das interfaces de inúmeros produtos envolve não apenas o aspecto da tradução do idioma original, mas sobretudo a consideração dos requisitos e necessidades de outros países e culturas [Niel93c, delG96, Mand97, Port97, Dix98]. A sofisticação tecnológica crescente em diversos países e o intercâmbio mundial de tecnologias, resultante desse crescimento, implicam uma necessidade maior de observação de aspectos internacionais no projeto e desenvolvimento de interfaces homem-máquina.

Do ponto de vista do fabricante, o aumento do número de vendas em outros países deverá significar que as vendas futuras dependerão não apenas dos aspectos de usabilidade considerados localmente, mas também estarão intimamente vinculadas a aspectos internacionais de usabilidade (e.g., questões idiomáticas e culturais). Por outro lado, do ponto de vista do usuário, constata-se que, para o contingente significativo de usuários de produtos interativos que lidam ou lidarão com interfaces originalmente projetadas em países estrangeiros, a usabilidade dependerá da conscientização sobre questões relativas ao projeto de interfaces usuário-computador para aplicação internacional [Niel93c, delG96].

Infelizmente, constata-se, mesmo em nível do país de origem, o lançamento de produtos interativos de naturezas as mais diversas que apresentam diferentes classes de problemas de usabilidade [Jord96a, Mayh99]. Várias indústrias e empresas têm surgido com o propósito de dar suporte à operação de produtos e sistemas baseados em computadores, equipamentos eletrônicos e, até mesmo, eletrodomésticos adquiridos por uma infinidade de indivíduos que não conseguem usá-los corretamente ou não são capazes de usufruir de todas as potencialidades oferecidas [Rubi94].

Há diversos exemplos de produtos de suporte desenvolvidos especialmente para realizar, de modo menos complicado (conforme apregoam seus fabricantes), funcionalidades oferecidas por produtos e sistemas baseados em computadores, equipamentos eletrônicos e eletrodomésticos de uso geral. Outra categoria de produtos de suporte que tem invadido o mercado de consumo são os vídeos de treinamento e aplicativos multimídia destinados ao ensino do uso correto de produtos específicos.

Ironicamente, tanto os equipamentos eletrônicos quanto os aplicativos de *software* para os quais produtos de suporte tiveram que ser desenvolvidos são, em geral, adquiridos antes da aquisição do "suporte de uso", seja porque os usuários pressupõem que uma de suas características é a *facilidade de uso* ou porque tais produtos de suporte ainda não foram desenvolvidos.

Tal situação, oportuna para diversos empresários, que lucram às custas de falhas existentes no processo de comunicação entre usuários e produtos interativos, só reforça o fato de que tornou-se comum o lançamento de produtos dotados de inúmeras funcionalidades, porém difíceis de usar por um grande contingente de consumidores.

Evidentemente, a oportunidade empresarial proporcionada por tal situação não é exclusiva aos fabricantes de produtos de suporte, já que os fabricantes originais também têm chances similares de elevarem seus lucros com o desenvolvimento de produtos genuinamente fáceis de aprender e usar logo após terem sido adquiridos. O que parece promissor em todo esse cenário é o crescente interesse e ênfase dados à usabilidade por diversos desenvolvedores de *software*.

Tal estágio de conscientização tem implicado necessidades que vêm possibilitando o crescimento do número de sociedades e associações destinadas à pesquisa de questões relativas à usabilidade, à elaboração de diretivas nacionais e internacionais para a usabilidade de produtos e a divulgação da importância desta característica para um mercado consumidor sempre crescente. Este fato tem alterado significativamente o modo como os produtos são projetados, desenvolvidos e avaliados.

Embora o campo de estudo se tenha restringido, inicialmente, a especialistas em fatores humanos ou ergonomistas, a demanda atual por produtos usáveis tem ultrapassado o número de profissionais especializados disponíveis. Este fato tem feito com que muitos desenvolvedores de produto, engenheiros, projetistas de sistemas, comunicadores técnicos e especialistas em treinamento e *marketing* passem a assumir responsabilidades diretas por questões de usabilidade em diversas organizações do mundo inteiro, que lhes exigem, em diferentes níveis e contextos, a execução de tarefas relativas à engenharia da usabilidade, para as quais não foram suficiente ou devidamente capacitados. Daí, a importância de estudos e a concepção de procedimentos que possam servir de referência e auxílio para profissionais atuantes em diferentes situações e contextos de trabalho que envolvam questões de usabilidade.

Outra questão que motiva e, ao mesmo tempo, justifica a realização de pesquisas na área de avaliação da usabilidade de produtos interativos diz respeito à tríade tecnológica formada pelos institutos de pesquisa, instituições acadêmicas e corporações de desenvolvimento de produtos. Em geral, pesquisadores, instituições de ensino e desenvolvedores se revezam no tocante ao estudo e compreensão da eficiência, eficácia, aplicabilidade, adaptabilidade e facilidade de aprendizado de diferentes estratégias de avaliação da usabilidade de produtos interativos usuário-computador.

No processo de concepção de estratégias avaliatórias, os pesquisadores carecem de casos de estudo, assim como de resultados advindos de experimentos conduzidos a partir desses casos, que lhes possibilitem otimizar e expandir da forma mais adequada possível os procedimentos metodológicos desenvolvidos [John96a]. As instituições de ensino, por sua vez, necessitam de informações que lhes permitam atualizar os conteúdos programáticos da diversidade (sempre crescente) de cursos na área de processos interativos usuário-computador, ao mesmo tempo que lhes proporcionem material para novas linhas de investigação científica. Por fim, os desenvolvedores de produtos carecem de informações sobre quais procedimentos metodológicos se adequam melhor aos recursos humanos, físicos e materiais de que dispõem, assim como a seus planos de trabalho e cronogramas de atividades.

No sentido de oferecer subsídios tanto às instituições de ensino quanto às organizações de desenvolvimento de produtos interativos de *hardware/software*, alguns autores (e.g., Jeffries *et al.* [Jeff91], Desurvire *et al.* [Desu92], Desurvire [Desu94], Karat [Kara94a], Karat *et al.* [Kara92]) têm produzido trabalhos destinados à comparação formal de diferentes estratégias de avaliação, evidenciando o potencial e as limitações de tais estratégias a partir de indicadores de desempenho e confiabilidade (número e tipo de problemas de usabilidade detectados, previsão de reações comportamentais do usuário, tempo de execução do procedimento metodológico, etc.) e/ou de relações de custos/benefícios.

Entretanto, muitos dos estudos dessa natureza não oferecem informações sobre o que os agentes do processo realmente fazem ao adotarem uma determinada técnica de avaliação ou uma

dada combinação de técnicas [Quei97b]. A carência de informações mais detalhadas do processo talvez seja o maior empecilho para a compreensão de como a técnica ou o conjunto de técnicas adotado permitirá ao avaliador identificar problemas de usabilidade, independentemente de sua capacidade. Além do mais, impossibilita os desenvolvedores de produtos de saberem o que esperar ao adotarem uma nova técnica de avaliação. Este inconveniente também dificulta um retorno efetivo de informações aos pesquisadores responsáveis pela concepção das técnicas, a fim de poderem otimizá-las [John96a].

A avaliação de um processo ou de um produto consiste usualmente de uma adaptação metodológica das técnicas consideradas às condições espaço-temporais, humanas e materiais disponíveis. Eis porque o inconveniente acima ressaltado fatalmente se estenderá à abordagem metodológica global, na qual as técnicas de avaliação propriamente ditas se inserem. Esta situação se assemelha a um estudo da usabilidade de uma interface cujos resultados são apresentados puramente sob a forma de um conjunto de tabelas contendo indicadores objetivos¹⁵ ou de uma listagem de indicadores subjetivos¹⁶. Na realidade, o resultado que possibilitaria a avaliação das técnicas de usabilidade adotadas deveria ser uma espécie de protocolo escrito relativo à abordagem metodológica considerada [John96a].

Quando se descreve uma descoberta em nível da pesquisa pura, não se relata apenas o que foi apreendido, mas também se diz como as respostas foram obtidas e se apresenta os dados coletados e o modo de processamento desses dados, procedimento que se estende à pesquisa aplicada [Glas97]. Estudos de casos se afiguram vantajosos com relação a inspeções, ensaios e outras estratégias de pesquisa à medida que conseguem responder os como e os porquês envolvidos no contexto de um conjunto de eventos sobre os quais o investigador possui pouco ou nenhum controle [Yin94].

Esta assertiva parece aplicável à avaliação de abordagens metodológicas adotadas no domínio dos processos interativos usuário-computador. Nestas circunstâncias, em que não é exercido nenhum controle potencial sobre como o avaliador apreende ou emprega uma determinada técnica (ou conjunto de técnicas), no contexto de um enfoque avaliatório, o questionamento versa sobre como tal técnica poderá ser empregadas para evidenciar problemas de usabilidade e sobre o porquê daquela técnica se adequar a algumas situações e não a outras [John96a]. Além disso, há questões relativas às potencialidades e limitações intrínsecas e extrínsecas ao enfoque avaliatório que merecem ser melhor investigadas e compreendidas, constituindo outro fator estimulante para a pesquisa.

1.4 Objeto da Pesquisa

O aprofundamento das características, potencialidades e limitações das técnicas e enfoques avaliatórios documentados na literatura clássica da área e discutidos apenas em linhas gerais nas sub-seções anteriores demanda uma revisão mais detalhada, constituindo um dos objetos deste trabalho. Assim, particularidades de aspectos relativos às técnicas e enfoques avaliatórios tratados

¹⁵ Tais como estimadores de desempenho, e.g., número de ações incorretas, número de consultas aos recursos de ajuda *online*, número de erros repetidos.

¹⁶ Tais como indicadores de satisfação, e.g., a opinião do usuário sobre a facilidade de localização de informações na documentação ou de aprendizado de novas potencialidades do produto.

- numerar fontes discutidas
- será 1.3
e 1.3 será
- retirar 7
aspas

neste capítulo com uma conotação introdutória serão explorados, de forma mais aprofundada, no próximo capítulo.

Além disso, a discussão introduzida nas sub-seções precedentes suscita principalmente uma investigação de como as componentes centrais de cada enfoque clássico influenciam o(s) atributos-chaves almejados para o produto-alvo da avaliação. Tal investigação constitui o objeto-chave desta pesquisa, que será tratado a partir do Capítulo 3 deste trabalho.

Em linhas gerais, o presente trabalho objetiva a confrontação de três enfoques avaliatórios clássicos, centrados na usabilidade de produtos interativos usuário-computador – *mensuração do desempenho*, *mensuração da satisfação* e *inspeção de padrões* – com o propósito de investigar a influência do atributo central¹⁷ de cada enfoque considerado sobre a usabilidade da interface de um aplicativo de *software* de uso corrente no âmbito da Engenharia Elétrica.

Neste contexto, também deverão ser investigados aspectos das seguintes relações de compromisso na produtividade associada ao uso: (i) *desempenho x satisfação subjetiva*; (ii) *desempenho x conformação a um padrão*; e (iii) *satisfação subjetiva x conformação a um padrão*.

Para tanto, a investigação efetuada ao longo da execução da pesquisa fundamentou-se nas seguintes hipóteses:

- H1: A usabilidade será tanto maior quanto maior for o desempenho associado ao uso do produto.
- H2: A usabilidade será tanto maior quanto maior for a satisfação subjetiva associada ao uso do produto.
- H3: A usabilidade será tanto maior quanto maior for o grau de adequação do produto a um padrão internacional.
- H4: O desempenho será tanto maior quanto maior for a satisfação subjetiva associada ao uso do processo/produto.
- H5: A satisfação subjetiva será tanto maior quanto maior for o desempenho associado ao uso do processo/produto considerado.
- H6: O desempenho será tanto maior quanto mais fundamentado em um padrão (internacional) for o desenvolvimento do produto considerado.
- H7: A satisfação subjetiva será tanto maior quanto mais fundamentado em um padrão (internacional) for o desenvolvimento do produto considerado.

¹⁷ *Desempenho* para a *mensuração do desempenho*, *satisfação subjetiva do usuário* para a *mensuração da satisfação* e *grau de adequação a um padrão* para o enfoque *inspeção de padrões*.

1.5 Objetivos da Pesquisa

1.5.1 Geral

Desenvolver um estudo analítico contextualizado da importância do desempenho de uso, da satisfação subjetiva associada ao uso e do grau de adequação de um produto a um padrão internacional para a usabilidade de tal produto, focalizando a influência de cada um desses atributos de usabilidade sobre os demais.

1.5.2 Específicos

- Investigar a influência dos compromissos *desempenho x satisfação subjetiva*, *desempenho x conformação a um padrão* e *satisfação subjetiva x conformação a um padrão* na produtividade associada ao uso;
- Disponibilizar um procedimento metodológico contextual constituído de uma abordagem avaliatória híbrida de produtos interativos, fundamentada tanto *na interação usuário-produto* quanto *na inspeção do produto por especialistas*, passível de adaptação e aplicação a outros contextos.

1.6 Encaminhamento Metodológico

É conveniente considerar que esta seção visa tão somente apresentar uma visão sinóptica do encaminhamento metodológico dado a esta pesquisa. O detalhamento dos enfoques adotados, assim como a discussão de aspectos específicos a cada um deles, ocorrerá nos Capítulos 5 e 6.

Em linhas gerais, a metodologia de trabalho considerada nesta proposta de pesquisa foi concebida a partir dos objetivos a serem atingidos, assim como de uma série de condições materiais, humanas e econômicas restritivas, dentre as quais podem ser citadas: a *escassez de avaliadores e de recursos instrumentais*, a *escassez de recursos financeiros* e o *número reduzido de usuários de teste disponíveis para recrutamento em caráter voluntário*.

O Quadro 1 apresenta, de forma generalizada, as cinco etapas que compõem a parte experimental desta pesquisa. Para cada etapa, são também discriminados os objetivos a atingir e as estratégias a adotar a fim de atingi-los. Embora se mencione cinco etapas, observa-se que o Quadro 1 contém uma etapa rotulada TREINAMENTO DO UNIVERSO AMOSTRAL, considerada anteriormente no contexto do processo avaliatório de um sistema de informações geográficas (SIG) por Queiroz e Turnell [Quei98a, Quei98b], esta etapa foi suprimida do contexto desta pesquisa, por se afigurar dispensável aos propósitos do processo avaliatório, conforme será mais bem observado no Capítulo 5. Sua inclusão no Quadro 1 deve-se ao fato de diversos contextos avaliatórios implicarem a inclusão de uma etapa de treinamento que proporcione a capacitação de um grupo de usuários de teste no uso do produto a ser avaliado ou lhes permita adquirir familiaridade com aquele produto. Isto ocorre sobretudo em contextos avaliatórios que envolvem usuários de teste principiantes sem nenhuma experiência prévia com o produto.

Quadro 1 – Generalização da abordagem metodológica

ETAPA	SUBETAPAS
PLANEJAMENTO DOS EXPERIMENTOS AVALIATÓRIOS	<ul style="list-style-type: none"> ▪ Caracterização do alvo de estudo ▪ Definição das metas e interesses relativos ao processo avaliatório ▪ Definição do perfil de usuário para caracterização do universo amostral ▪ Mapeamento dos usuários de teste potenciais disponíveis ▪ Definição do modo de recrutamento dos usuários de teste ▪ Decisão do número de participantes do ensaio avaliatório ▪ Seleção das técnicas de avaliação ▪ Definição de indicadores subjetivos e/ou objetivos
TREINAMENTO DO UNIVERSO AMOSTRAL ¹⁸	<ul style="list-style-type: none"> ▪ Familiarização dos usuários de teste com a ferramenta a ser testada ▪ Capacitação do universo amostral de teste no aplicativo a ser avaliado
ELABORAÇÃO DO MATERIAL DO ENSAIO	<ul style="list-style-type: none"> ▪ Definição e estruturação das atividades associadas à avaliação do produto ▪ Definição dos recursos necessários à execução de cada tarefa ▪ Elaboração da ficha cadastral e do documento de aceitação das condições de teste ▪ Elaboração do material necessário à condução do processo avaliatório ▪ Discussão das abordagens de avaliação da usabilidade do produto ▪ Modelagem da tarefa e da interação ▪ Elaboração dos roteiros da tarefa ▪ Validação do material elaborado
CONDUÇÃO DO ENSAIO E COLETA DE DADOS	<ul style="list-style-type: none"> ▪ Aplicação das estratégias avaliatórias pré-definidas ▪ Registro dos indicadores pré-definidos
TABULAÇÃO E ANÁLISE DOS DADOS	<ul style="list-style-type: none"> ▪ Triagem preliminar dos dados coletados ▪ Triangulação dos dados coletados ▪ Tabulação e síntese dos dados coletados ▪ Organização dos problemas listados
APRESENTAÇÃO DOS RESULTADOS	<ul style="list-style-type: none"> ▪ Definição do modo de divulgação dos resultados ▪ Priorização dos dados apresentados de acordo com os usuários finais dos resultados ▪ Elaboração do documento final

A adoção dos enfoques avaliatórios mencionados na Seção 1.4 (extensíveis e adaptáveis a outros contextos avaliatórios), com base considerações feitas nas três seções iniciais deste capítulo, carece de uma reflexão sobre os mais variados aspectos característicos de cada enfoque (e.g. viabilidade dos procedimentos associados a cada uma de suas etapas, clareza dos objetivos a atingir ao longo de cada etapa, grau de detalhamento da descrição, facilidade de aprendizado e execução do enfoque, consistência dos resultados face às técnicas de avaliação adotadas, respaldo dos resultados em outras iniciativas de natureza similar).

Análises reflexivas desta natureza atuam usualmente como uma espécie de avaliação do processo de avaliação que se pretende adotar, podendo serem fundamentadas pela eleição de um conjunto de critérios que possibilitem a verificação da confiabilidade das informações referentes ao potencial e às limitações apontadas pelos proponentes das técnicas de avaliação consideradas.

Os critérios de confiabilidade considerados neste trabalho para a análise das abordagens avaliatórias do produto considerado foram: (i) o grau de detalhamento da descrição dos procedimentos avaliatórios; (ii) a dimensão da base de dados coletados; (iii) a fundamentação dos procedimentos em outros ensaios de natureza similar; e (iv) a consistência dos resultados obtidos face às técnicas de avaliação adotadas. Reforçando o que já foi mencionado no segundo parágrafo

¹⁸ Esta é uma etapa opcional no contexto global da abordagem metodológica considerada.

desta seção, a aplicação dos critérios acima mencionados ao contexto deste trabalho poderá ser verificada a partir da leitura dos Capítulos 7 e 8.

1.7 Organização do Documento

Este documento foi subdividido em oito capítulos. O presente capítulo visou situar o trabalho no domínio mais amplo de estudo dos processos avaliatórios de produtos interativos usuário-computador, ao mesmo tempo que focalizou o objeto de estudo da proposta de pesquisa. Além disto, esta seção complementa a visualização da estrutura dos demais capítulos deste documento, à medida que apresenta nos parágrafos seguintes uma descrição sumária de seus conteúdos.

O Capítulo 2 apresenta uma revisão da literatura técnica da área de concentração do trabalho, descrevendo e discutindo quatro categorias de técnicas consagradas de avaliação da usabilidade de produtos interativos usuário-computador, com o intuito de apresentar o domínio de conhecimento que encerra o objeto da pesquisa proposta. Neste capítulo, também se encontra uma revisão sobre métodos empíricos de avaliação, além de outros métodos resultantes de estudos experimentais que envolvem *comparações* entre técnicas e/ou estilos de interação.

O Capítulo 3 retoma a discussão de aspectos relativos à avaliação da usabilidade não explorados nos capítulos anteriores, ressaltando questões pertinentes ao contexto dos enfoques avaliatórios consagrados na literatura da área, especialmente aquelas concernentes à mensuração da usabilidade, onde é retomada e expandida a discussão de aspectos referentes à seleção de um enfoque avaliatório, segundo diferentes perspectivas teóricas típicas de processos interativos usuário-computador, e revisadas características dos diferentes enfoques discutidos no capítulo anterior. Neste capítulo também são feitas considerações sobre o caráter complementar dos enfoques avaliatórios quantitativo e qualitativo.

O Capítulo 4 discute a importância do uso de padrões em processos avaliatórios, ressaltando benefícios e limitações inerentes a esta estratégia de avaliação. Nele são comentados, em linhas gerais, padrões vigentes de usabilidade e, em maiores detalhes, o padrão ISO 9241, empregado na inspeção conduzida no contexto desta pesquisa. No entanto, a seção mais relevante deste capítulo descreve detalhadamente a metodologia adotada na inspeção do produto-alvo conforme recomendações contidas nas partes 11, 14, 15 e 16 do padrão ISO 9241.

Os Capítulos 5 e 6 apresentam uma estruturação coerente com aquela adotada no Capítulo 4, discutindo, respectivamente, a importância, os benefícios e as limitações da avaliação do desempenho e da satisfação do usuário, e descrevendo detalhadamente os procedimentos metodológicos adotados neste trabalho.

O Capítulo 7 apresenta e discute os resultados obtidos no contexto da pesquisa a partir dos procedimentos metodológicos apresentados nos três capítulos anteriores, além de confrontá-los, objetivando verificar as hipóteses que fundamentam a tese apresentada na Seção 1.4.

Finalmente, o Capítulo 8 apresenta as conclusões advindas dos resultados obtidos com a aplicação da metodologia descrita nos Capítulos 4, 5 e 6, evidenciando as principais contribuições desta pesquisa para a área de avaliação de interfaces usuário-computador e sugerindo temas para trabalhos futuros.

2

CONSIDERAÇÕES SOBRE QUALIDADE E USABILIDADE DE SOFTWARE

"Some designers cannot understand the apparent stupidity of users who have difficulty accessing the brilliant functions which, at great cost, have been cleverly built into the software. HCI process method producers need to recognise that they are now in the same situation, with those stupid system designers failing to use their sophisticated usability design guides and procedures. So how do we set about taking our own medicine?"

[Brian Shackel]

2	Considerações sobre Qualidade e Usabilidade de Software	23
2.1	Qualidade de Software e Satisfação do Usuário	25
2.1.1	Qualidade e Qualidade de Software	26
2.3.2	Facetas da Qualidade de Software.....	34
2.3.2	Qualidade de Software e Satisfação do Usuário	39
2.2	Engenharia, Engenharia de Software e Engenharia da Usabilidade	41
2.2.1	Engenharia.....	41
2.2.2	Engenharia de Software.....	43
2.2.3	Engenharia da Usabilidade.....	45
2.3	Usabilidade de Software	47
2.3.1	Visão de Shackel.....	53
2.3.2	Visão de Nielsen.....	55
2.3.3	Visão da ISO (<i>International Standards Organization</i>)	57

O presente capítulo discute aspectos relativos à qualidade e usabilidade de produtos de *software* e às técnicas de avaliação de processos interativos usuário-computador, suas potencialidades e limitações.

A seção 2.1 - **Qualidade de Software e Satisfação do Usuário** – se inicia com definições do termo *qualidade* que servem como ponto de partida para uma discussão sobre a *qualidade de software*. São apresentadas definições de diversos autores consagrados, assim como modelos para a mensuração da qualidade de produtos de *software*. Finalmente, a seção é encerrada com comentários sobre a relação existente entre a qualidade de *software* e a satisfação do usuário.

A seção 2.2 (**Engenharia, Engenharia de Software e Engenharia da Usabilidade**) se inicia com um estudo etimológico da palavra *engenharia* e apresenta definições de diversos autores para os termos *engenharia de software* e *engenharia da usabilidade*, objetivando justificar o emprego deste último ao longo deste documento.

A última seção deste capítulo (**Usabilidade de Software**) retoma a discussão conduzida na primeira seção, buscando evidenciar o elo existente entre a qualidade e a usabilidade de *software*. Adicionalmente, apresenta considerações sobre usabilidade, discute três visões clássicas da usabilidade enquanto atributo mensurável de produtos e direciona a discussão para o âmbito da avaliação de interfaces usuário-computador, sendo encerrada com uma breve menção às técnicas concebidas para este propósito.

2.1 Qualidade de Software e Satisfação do Usuário

É inegável que, no último quarto de século, as tecnologias de *hardware* e *software* evoluíram a passos largos, estimulando o uso de sistemas baseados em computadores nas mais diferentes atividades do cotidiano e instituindo uma nova ordem em nível mundial. Após décadas de concentração de esforços direcionados para o desenvolvimento e o aprimoramento de tecnologias computacionais, as tendências atuais da engenharia de sistemas apontam para estratégias de concepção e desenvolvimento de projetos que visam cada vez mais a solução de problemas *reais* do usuário.

Também é inegável que os sistemas de *software* tornaram-se praticamente onipresentes no cotidiano da vida contemporânea, uma vez que a maioria dos equipamentos elétricos incluem algum tipo de *software* [Somm97]. Nas duas últimas décadas, em especial, o *software* vem assumindo cada vez mais o papel do *hardware* como componente de planejamento mais difícil, menos provável de obter sucesso tanto em termos de cumprimento de prazos quanto do orçamento de custos e mais arriscado de se administrar [Press95, Somm97]. Em suma, demandando mais atenção e esforços. Em geral, constata-se que os custos envolvidos em processos de desenvolvimento de *software* não estão associados a características do produto relacionadas com a metodologia adotada para implementá-lo, nem tampouco com a confiança dos resultados que tal metodologia proporciona. Os custos estão associados, sobretudo, a quão adequadamente o *software* é projetado, codificado e documentado, no que diz respeito à manutenção, portabilidade, interoperabilidade e atualização, dentre outros aspectos [McCa79].

Os recursos investidos por grande parte das corporações de desenvolvimento de produtos de *software* são cada vez mais vultuosos. Devido ao aumento da complexidade dos produtos, dos requisitos de qualidade exigidos e das demandas de mercado, muitos dos problemas encontrados são específicos ao processo de desenvolvimento de *software*, dentre os quais podem ser citados dificuldades de planejamento, ausência ou precariedade de bons programas de avaliação da qualidade dos produtos desenvolvidos, não cumprimento de prazos e, até mesmo e sobretudo, falta de estruturação das atividades de desenvolvimento e avaliação dos produtos. Esta é uma das principais razões que têm impulsionado companhias e instituições diversas do mundo inteiro a continuarem investindo em recursos físicos, materiais, temporais e econômicos em processos de otimização do desenvolvimento das aplicações de *software* que produzem e/ou utilizam.

Problemas de desenvolvimento enquadrados neste contexto compõem um quadro usualmente rotulado por diversos pesquisadores da área (e.g., Gibbs [Gibb94], Sommerville [Somm97]) como *crise de software* (*software crisis*). Uma análise da situação atual do mercado de desenvolvimento de *software* conduz à constatação de que tais problemas vêm recebendo uma atenção sempre crescente e sendo tratados cada vez mais efetivamente no contexto da Otimização do Processo de Desenvolvimento de Software (SPI¹⁹). Como consequência, diversas abordagens, envolvendo modelos, métodos e métricas variadas têm sido concebidas e validadas por pesquisadores e equipes de desenvolvimento de *software*.

Em geral, estas abordagens costumam ser enquadradas por alguns autores (e.g., Thomas e McGarry [Thom94], Pressman [Press95], Sommerville [Somm97], Solingen e Berghout [Sol97])

¹⁹ *Software Process Improvement*.

em uma de duas grandes categorias, assim discriminadas:

(i) abordagens *top-down*, tais como *SPICE* [Dor193], *BOOTSTRAP* [Kuva94] e *CMM* [Paul95], fundamentadas na determinação e análise do esforço global de desenvolvimento de *software*, segundo características pré-definidas; e

(ii) abordagens *bottom-up*, tais como *AMI* [AMI92], *GQM* [Basi94a] e *QIP* [Basi94b], que se fundamentam essencialmente na mensuração de componentes do processo de desenvolvimento, sem características pré-definidas como fontes de informação básicas.

Em essência, a linha de abordagem *top-down* aplica um modelo normativo considerado o melhor modo de desenvolvimento de *software*, que permite a análise da maturidade de uma empresa de desenvolvimento de *software* e a proposição de otimizações relevantes ao contexto do processo de desenvolvimento [Hump89]. Em contrapartida, a linha de abordagem *bottom-up* argumenta que o campo da Engenharia de *Software* ainda é imaturo, tornando impossível a definição de tal modelo normativo para desenvolvimento de *software* [Gibb94]. Portanto, adquire-se conhecimento empírico através da mensuração de componentes do processo de desenvolvimento de *software*, visando o aprimoramento da compreensão do referido processo em contextos específicos. Nos últimos anos, ambas as linhas de abordagem vêm sendo aplicadas satisfatoriamente em diversos contextos práticos [Somm97].

Sendo um processo através do qual números ou símbolos são associados a atributos de entidades do mundo real, possibilitando sua descrição de acordo com regras claramente definidas [Fent96], a mensuração constitui um auxílio importante em abordagens *bottom-up* de otimização de processos de desenvolvimento de *software*, que sempre incluem algum tipo de plano de mensuração. Nesse contexto, a qualidade do produto representa um papel de destaque, exercendo uma forte influência no cotidiano de pesquisadores, projetistas, avaliadores e usuários, pois dela depende a viabilidade das atividades que estes realizam com o auxílio do produto [Kitc96]. Atualmente, a motivação para a perseguição da qualidade se afigura evidente, graças à sua introdução em todos os níveis da produção de bens por técnicos, à sua popularização pela mídia e à sua elevação de *status* como critério decisivo na aquisição de bens pelos consumidores.

2.1.1 Qualidade e Qualidade de Software

A busca da qualidade de produtos de *software* implica a consideração de uma série de fatores os mais variados, abrangendo questões que vão desde a tomada de decisões por parte dos empresários até a seleção de métricas mais apropriadas, passando por detalhes de cunho temporal e econômico concernentes aos investimentos a serem feitos pelas empresas de desenvolvimento. Um ponto de partida apropriado para uma discussão do assunto pode ser o questionamento do termo *qualidade*.

Sem dúvida, a definição do termo *qualidade* constitui um tema controverso. Uma discussão interessante do significado do termo pode ser encontrada em Kitchenham e Walker [Kitc86]. Hyatt e Rosenberg [Hyat96] comentam que embora todos concordem que a *qualidade* é importante, mas ninguém concorda com o que *qualidade* significa. Kitchenham [Kitc96] afirma que *qualidade* é difícil de definir e impossível de medir, porém fácil de reconhecer. Gillies [Gill92] acrescenta que a *qualidade* é transparente quando presente, porém facilmente reconhecida em sua ausência.

Por outro lado, as definições dos dicionários são excessivamente vagas ou sujeitas a uma grande diversidade de interpretações para servirem de respaldo para uma discussão. O *Novo Dicionário da Língua Portuguesa* [Holl86], por exemplo, define *qualidade* como "1. Propriedade, atributo ou condição das coisas ou das pessoas capaz de distingui-las das outras e de lhes determinar a natureza. 2. Numa escala de valores, qualidade (1) que permite avaliar e, conseqüentemente, aprovar, aceitar ou recusar qualquer coisa."

O *Oxford Advanced Learner's Dictionary* [Horn90], por sua vez, define o termo equivalente inglês *quality* como "1 (a) grau de excelência ou mérito (*degree of goodness or worth*). (b) excelência global (*general excellence*). 2 (a) atributo, característica (*attribute, characteristic*). (b) aspecto especial ou distintivo (*special or distinguishing feature*)."

Percepções de *qualidade* por diferentes domínios do conhecimento humano, incluindo filosofia, economia e *marketing*, foram examinadas e relatadas por Garvin [Garv84, Garv88], que as encarou como diferentes facetas de um conceito complexo. O autor descreveu as facetas do termo a partir de cinco perspectivas diferentes, a saber:

(i) *Transcendental*, baseada na descrição platônica do *ideal* e no conceito aristotélico de *forma*, este ponto de vista encara a *qualidade* como algo que pode ser reconhecido, mas não definido;

(ii) *Do usuário*, embasada no compromisso entre as suas necessidades e as características do produto (altamente personalizada e inerente ao contexto de modelagem de desempenho e confiabilidade, visto que se avalia comportamentos de produtos com relação à funcionalidade esperada e aos padrões de uso);

(iii) *Do produto*, fundamentada em características inerentes ao produto, assumindo que a mensuração e o controle de suas propriedades intrínsecas (denominadas *indicadores de qualidade internos*) resultam na otimização de seu comportamento externo (*qualidade em uso*);

(iv) *Do fabricante*, focalizada na *qualidade* do produto durante o seu desenvolvimento e após a entrega (ou seja, dependente da especificação);

(v) *Baseada no custo*, respaldada no montante que o consumidor se dispõe a pagar pelo produto, o que implica a avaliação e revisão dos requisitos de produção à luz de relações de custos e benefícios.

Portanto, a definição de *qualidade* está intimamente atrelada à perspectiva que se adote para visualizá-la. Mensurá-la significa estabelecer referências, prognosticá-la adequadamente e monitorar processos que impliquem sua otimização, ou seja, vinculá-la à perspectiva sob a qual é encarada. Diferentes perspectivas implicam a consideração de diferentes aspectos que exercem influência sobre a qualidade de um produto e, por conseguinte, a concepção de diferentes modelos de *qualidade*. O Quadro 2, adaptado de uma das páginas do site da *Tantara Management Services* [TMS00], complementa a discussão sobre *qualidade*, sintetizando aspectos relativos aos pontos de vista de seis pesquisadores clássicos da área, a saber: Deming [Demi86, Watt86], Garvin [Garv88, Flow90], Taguchi [*in* Roy90], Feigenbaum [Feig91], Crosby [Cros92] e Juran [Jura95].

Quadro 2 – Visões de qualidade.

QUALIDADE	AUTOR					
	Deming [Walt86]	Garvin [Garv88]	Tagushi [Roy90]	Feigenbaum [Feig91]	Crosby [Cros92]	Juran [Jura95]
Orientação básica	À técnica	À estratégia	À técnica	Ao sistema	À motivação	Ao processo
Definição	Sistemas sem falhas	Oportunidade competitiva	Exigências do cliente quanto ao desempenho	O que o cliente diz que é	Adequação às exigências	† Adequação ao uso † Ausência de problemas
Responsável	Gerente de desenvolvimento	Gerente de desenvolvimento	Engenheiros	Todo o mundo	Gerente de desenvolvimento	Gerente de desenvolvimento
Importância das exigências do cliente como padrão	Muito importante	Muito importante	Muito importante	Muito importante	Muito importante	† Muito importante † Clientes presentes em cada etapa do ciclo de vida do produto
Meta	† Atendimento/superação das necessidades do cliente † Otimização ininterrupta do produto	† Satisfação do cliente † Otimização ininterrupta do produto	† Atendimento das necessidades do cliente † Otimização ininterrupta do produto	† Atendimento das necessidades do cliente † Otimização ininterrupta do produto	† Otimização ininterrupta do produto † Ausência total de falhas do produto	† Satisfação do cliente † Otimização ininterrupta do produto
Métodos para atingi-la	† Métodos estatísticos † Consonância nas metas e propósitos † Otimização ininterrupta † Ação cooperativa entre funções	† Identificação de nichos de qualidade	† Métodos estatísticos e.g Função de Perda † Supressão de variações nas características de projeto e do "ruído" através de processos e projetos robustos	† Controle da qualidade total (TQC) † Processo direcionado para a excelência do produto, não para suas falhas	† Abordagem de 14 pontos	† Custos de qualidade † Trilogia de qualidade † Planejamento † Controle † Otimização
Elementos-chaves da implementação	† Programa de 14 etapas;	† 8 dimensões de qualidade de produto • desempenho • características • confiabilidade • adequação • durabilidade • serviceability • estética • qualidade percebida	† Projeto estatístico de ensaios; equipes de qualidade	† Métodos estatísticos e de engenharia em todos os setores de desenvolvimento da empresa	† Programa de 14 etapas † Custo de qualidade † Emprego da "grade de maturidade" (<i>maturity grid</i>) para a gestão da qualidade	† Programas de inovação † Conselho de qualidade † Equipes de qualidade
Papel do treinamento	Muito importante para gerentes e funcionários	Importante, mas não claramente definido	Importante, mas não definido	Muito importante para gerentes e supervisores	Muito importante para gerentes e funcionários	Muito importante para gerentes e funcionários

Os Quadros 3, 4 e 5 apresentam, respectivamente o programa de 14 pontos sugerido por Deming [Demi86], as oito dimensões da qualidade apontadas por Garvin [Garv94] e o programa de 14 etapas sugerido por Crosby [Cros92] para a otimização da qualidade (referidos no Quadro 2).

Quadro 3 – O programa de 14 pontos de Deming [Demi86].

PROGRAMA DE 14 PONTOS DE DEMING	
1.	Tornar constante o propósito de otimização do produto e do serviço, objetivando a competitividade, a permanência no mercado e a oferta de empregos.
2.	Adotar a nova filosofia, pelo fato de se estar inserido em uma nova era econômica. (A administração ocidental deve despertar para os novos desafios, apre(e)nder suas responsabilidades e assumir a liderança através das mudanças.)
3.	Romper a dependência da obtenção da qualidade a partir da inspeção, eliminando a necessidade de inspeção respaldada em primeiro lugar na qualidade no produto.
4.	Encerrar a prática de compensação de custos à base da etiquetagem de preços. Ao invés disto, minimizar o custo total. Adquirir qualquer item necessário de um único fornecedor, estabelecendo uma relação de longo prazo baseada na lealdade e confiança.
5.	Otimizar Intermitentemente e sempre o sistema de produção e o serviço, a fim de otimizar a qualidade e a produtividade e, assim, progressivamente reduzir os custos.
6.	Instituir um programa de treinamento.
7.	Tornar efetiva a liderança (ver Ponto 12 e Capítulo 8 do <i>Out of the Crisis</i> , de Deming [Demi86]), instituindo como meta da supervisão o suporte que possibilite o exercício melhor do trabalho dos empregados, assim como das máquinas e dispositivos. A supervisão da administração carece de inspeção do mesmo modo que a supervisão dos empregados do setor de produção.
8.	Vencer os medos, de forma que todos possam trabalhar efetivamente para a companhia (ver Capítulo 3 do <i>Out of the Crisis</i> , de Deming [Demi86]).
9.	Romper as barreiras setoriais (equipes de pesquisas, projetos, vendas e produção), estimulando o trabalho em equipe, a fim de prevenir conjuntamente problemas na produção e do uso que possam ser encontrados no produto ou serviço.
10.	Eliminar <i>slogans</i> , exortações e metas que exijam da mão-de-obra a ausência de defeitos e novos níveis de produtividade. Tais exortações só criam relações de adversidade, quando a maior parte das causas da baixa qualidade e baixa produtividade é inerente ao sistema e, portanto, transcende o poder da mão-de-obra.
11.	a) Eliminar patamares-padrões de trabalho (cotas). b) Eliminar a gestão a partir do objetivo. Eliminar a gestão através de números e de metas numéricas.
12.	a) Remover barreiras que privem os empregados do direito de se sentirem orgulhosos de seu trabalho. Alterar a responsabilidade dos supervisores de meros números para qualidade. b) Remover barreiras que privem o pessoal da administração e da engenharia do direito de se sentirem orgulho de seu trabalho. Isto significa abolir a compensação anual por mérito e a gestão a partir do objetivo (ver Capítulo 3 do <i>Out of the Crisis</i> , de Deming [Demi86]).
13.	Instituir um programa vigoroso de educação e auto-aperfeiçoamento.
14.	Atribuir a todos o encargo de realizar a transformação. A transformação é trabalho de todos.

Quadro 4 – As oito dimensões da qualidade [Cros92].

AS OITO DIMENSÕES DA QUALIDADE DE GARVIN	
DIMENSÃO	DESCRIÇÃO
1. DESEMPENHO (<i>Performance</i>)	Características operacionais básicas do produto
2. FACILIDADES SUPLEMENTARES (<i>Features</i>)	Características que suplementam o funcionamento básico do produto
3. CONFIABILIDADE (<i>Reliability</i>)	Probabilidade de apresentação de falhas ou mau funcionamento em um determinado intervalo de tempo
4. CONFORMAÇÃO (<i>Conformance</i>)	Grau de adequação a padrões pré-estabelecidos
5. DURABILIDADE (<i>Durability</i>)	Tempo de vida até a substituição
6. PRESTAÇÃO DE SERVIÇOS (<i>Serviceability</i>)	Facilidade de assistência técnica, rapidez e competência dos serviços de manutenção
7. ESTÉTICA (<i>Aesthetics</i>)	Capacidade de satisfazer aos sentidos (aparência visual do produto e impressões auditiva, olfativa ou gustativa por ele causadas)
8. QUALIDADE PERCEBIDA (<i>Perceived quality</i>)	Inferências feitas a partir de aspectos tangíveis e intangíveis do produto (papel relevante e crítico da publicidade, da imagem e do rótulo do produto)

Quadro 5 – O programa de 14 etapas de Crosby [Cros92].

PROGRAMA DE 14 ETAPAS DE CROSBY
1. Evidenciar o comprometimento da administração com a qualidade.
2. Formar equipes de otimização da qualidade com representantes seniores de cada setor.
3. Mensurar processos com o propósito de determinar onde há problemas e/ou onde há potencialidade para a ocorrência de problemas.
4. Avaliar os custos da qualidade e explicar seu uso como ferramenta de gestão.
5. Estimular a preocupação individual de todos os empregados com a qualidade.
6. Agir no sentido de corrigir os problemas identificados a partir das etapas anteriores.
7. Estabelecer o monitoramento do progresso do processo de otimização (da qualidade).
8. Treinar supervisores a fim de que cumpram ativamente sua parte no programa de otimização da qualidade.
9. Instituir um "Dia sem Falhas" ²⁰ , a fim de que todos se conscientizem da mudança e reafirmem o comprometimento com o processo de gestão (da qualidade).
10. Encorajar os indivíduos a estabelecerem metas para a otimização em nível pessoal e grupal.
11. Encorajar os empregados a comunicarem aos gerentes os obstáculos de frontados para atingir suas metas de otimização.
12. Mostrar reconhecimento e apreciar aqueles que participam.
13. Estabelecer conselhos de qualidade que interajam regularmente.
14. Enfatizar que o programa de otimização da qualidade nunca acaba.

²⁰ Zero Defects Day, no original.

Dixon *et al.* [Dixo96] descrevem *qualidade* como uma associação evidente de características interrelacionadas que refletem adequadamente exigências específicas e implícitas de um produto, processo ou serviço que são relevantes em um dado contexto atual e sua circunvizinhança.

Segundo a visão da *European Organisation for Quality* [EOQ00], expressa através de um documento redigido durante a presidência finlandesa da União Européia, em 1999, *qualidade* é a relação entre requisitos e resultados reais, a diferença entre o que se espera e o que se consegue. O documento, que dedica duas seções inteiras discutindo aspectos relativos ao termo *qualidade*²¹, acrescenta que a qualidade é baseada em valores e expressa em escolhas, sendo um atributo valioso na distinção entre o bom e o ruim, o aceitável e o inaceitável, os sem iniciativa e os empreendedores. A fim de enfatizar as conotações do termo *valores*, o documento menciona que são vários os seus significados, embora apresente apenas os dois empregados na discussão, a saber: (i) idéias genéricas de preferências em situações onde se dispõe de duas ou mais linhas de ação; e (ii) no contexto de transações envolvendo produtos e serviços, atributo associado à percepção individual de utilidade e do grau de importância que cada indivíduo confere à transação. As duas conotações são sintetizadas na Fig. 1, adaptada do original [EOQ00], onde se ilustra do lado esquerdo o primeiro significado dado ao termo pela EOQ, enquanto o lado direito ilustra a segunda conotação que a EOQ dá ao termo.



Fig. 1 – Qualidade e valores, segundo a *European Organisation for Quality* [EOQ00].

Neste ponto, vale a pena questionar o que vem a ser, afinal de contas, *qualidade de software*. Na verdade, várias definições têm sido propostas para o termo na literatura da área. Em geral, a qualidade de um produto é tratada como uma de suas propriedades [Beva95a]. Assim, do ponto de vista do produto, a qualidade implica a identificação dos atributos que podem ser projetados ou avaliados, para que o produto seja otimizado. Na verdade, há quase duas décadas o termo *qualidade de software* (*software quality*) tem recebido uma atenção cada vez maior, tanto dos desenvolvedores quanto dos avaliadores e da comunidade usuária, tornando-se, por conseguinte, amplamente divulgado.

²¹ Seções 1 (The European Quality Vision - a background...) e 2 (Quality – philosophical and practice perspectives)

Segundo Solingen *et al.* [Soli99], quando se visa a qualidade de produtos de *software*, distingue-se duas abordagens possíveis, a saber: a *abordagem do processo* e a *abordagem do produto*. Os autores acrescentam que enquanto a primeira abordagem tenta otimizar a qualidade de produto indiretamente, a partir da otimização do processo, a segunda tenta fazê-lo diretamente. Todavia, em muitas circunstâncias, conotações vagas e equivocadas do termo, desvinculadas do *processo* e do *produto*, têm provocado uma certa desvirtuação da meta fundamental da indústria de *software*: o aprimoramento dos diversos processos envolvidos na geração de seus *produtos* finais [Drom96]. No que diz respeito aos usuários, Gentleman [Gent96] acrescenta que é surpreendente o número de indivíduos que ainda associa o termo *qualidade de software* simplesmente à ausência de erros.

Para alguns autores (e.g. Kitchenham e Pfleeger [Kitc96], Bevan [Beva97]), o contexto é um aspecto de grande significância [Kitc96, Gent96, Beva97, Bond98] para a definição do termo *qualidade de software*. O ponto de vista do qual o termo é focalizado representa outro aspecto relevante a ser considerado, tendo em vista que os resultados de julgamentos feitos por diferentes indivíduos sobre algo que usam são função de seus pontos de vista [Bond98], assim como de suas necessidades particulares. No âmbito da Engenharia de *Software*, não poderia ser diferente, tendo em vista que projetistas, usuários, fabricantes e avaliadores de ferramentas de *software* usualmente encaram o termo *qualidade* de ângulos distintos, ao aplicá-lo ao mesmo produto. São as chamadas "perspectivas da qualidade", mencionadas por diversos autores, e.g., Kusters *et al.* [Kust97] e Miller [Mill97].

Qualidade de software é freqüentemente definida em termos da adequação do produto aos propósitos segundo os quais foi desenvolvido. Entretanto, há de se considerar que diferentes usuários têm propósitos diferentes para o mesmo produto. Um usuário principiante esporádico provavelmente estará mais interessado na facilidade de aprendizagem e na tolerância a erros exibidas pelo produto do que na eficiência por este apresentada. Por outro lado, um gerente de redes que planeja incorporar o produto a algum sistema maior, estará mais interessado na capacidade de detecção e de recuperação de falhas que o produto apresenta do que na facilidade de instalação deste. Uma empresa de treinamento e manutenção do produto se preocupará com questões relativas à documentação técnica e a facilidade de teste do produto.

Tais considerações parecem conduzir à conclusão que qualidade de *software* não é absoluta, mas uma percepção dependente de quem a avalia. Além disso, a *qualidade de software* é multifacetada, sendo a importância de cada uma de suas facetas dependente do contexto de uso, até mesmo para o mesmo indivíduo.

Muitas referências clássicas em Engenharia de *Software* (e.g., Gilbert [Gilb83], Schach [Scha90], Hailstone [Hail91]) definem *qualidade de software* como a implementação correta da especificação. Tal definição pode ser adequada durante a fase de desenvolvimento de produto, embora se afigure inadequada quando utilizada para comparar produtos.

Kano *et al.* [Kano89] definem *qualidade de software* em três níveis, a saber: (i) *qualidade por excelência* (*excellent quality*), definida como qualquer característica do *software* que exceda o nível de desempenho esperado; (ii) *qualidade unidimensional*, definida como qualquer característica do *software* desejada por usuários específicos, porém não inteiramente necessária para todos as categorias de usuários; e (iii) *qualidade esperada*, definida como qualquer

característica necessária ao funcionamento normal do produto.

Denning [Denn92] focaliza o usuário ao definir *qualidade de software* e associa o termo ao grau de satisfação do usuário, afirmando que quanto maior for o grau de satisfação do usuário, mais provável será que este classifique o produto como confiável e de boa qualidade. Esta visão também é compartilhada pela ISO, através do padrão 8402 [ISO86], que define a *qualidade* de uma entidade (processo, produto ou organização) como a totalidade de características daquela entidade que lhe confere a capacidade de satisfazer as necessidades explícitas e implícitas. Tal definição estará fundamentada nas características do produto se as necessidades dos usuários forem bem definidas e comuns. Entretanto, partindo do princípio que grupos diferentes de usuários têm necessidades diferentes, então a qualidade do produto para cada grupo estará associada a características diferentes. Deste modo, a qualidade do produto, sob o ponto de vista do ISO 8402 torna-se dependente da percepção do usuário.

De um modo abrangente, as organizações nacionais e internacionais de padronização têm associado o termo *qualidade de produtos* (que dá margem à inclusão de produtos de *software*) ao atendimento de necessidades ou expectativas. O padrão IEEE 10.12-1990 [IEEE90b] define *qualidade* de dois modos, a saber: (1) grau segundo o qual um sistema, componente ou processo satisfaz as exigências especificadas; e (2) grau segundo o qual um sistema, componente ou processo satisfaz o cliente ou as necessidades ou expectativas do usuário.

Para o termo *qualidade de software*, a definição do padrão IEEE 1061-1992 [IEEE92] diz ser o grau segundo o qual o *software* possui uma combinação desejada de atributos. O padrão ISO 9000 [ISO87] diz respeito à garantia da *qualidade* de um produto sob o ponto de vista do preenchimento de um conjunto de requisitos pré-definidos. Segundo Bevan [Beva97], uma interpretação literal deste ponto de vista torna a qualidade de um produto dependente do profissional ou da equipe que especifica os requisitos. Dito de outro modo, um produto pode ser julgado como sendo de boa qualidade mesmo se a especificação dos requisitos não for apropriada a um dos contextos de uso para o qual foi desenvolvido.

Pressman [Pres95] não define *qualidade de software*. No entanto, o autor se refere ao termo como a conformação de todo produto de *software* profissionalmente desenvolvido a (i) requisitos funcionais e de desempenho explicitamente declarados, (ii) padrões de desenvolvimento claramente documentados e (iii) características implícitas almejadas (e.g., um elevado nível de portabilidade). O autor acrescenta que a *qualidade de software* é uma combinação complexa de fatores, também denominados *indicadores de qualidade*, que variam conforme as aplicações consideradas e os clientes que a solicitam.

Nos últimos 25 anos, têm-se concebido e proposto diversos modelos de *qualidade de software* que focalizam diferentes características do produto, às quais se tem associado uma série de fatores de qualidade. Tais modelos têm servido de suporte à compreensão de como os indicadores da *qualidade* de um produto se interrelacionam e às características consideradas, de modo a tornar praticável sua mensuração [Kitc96]. Dentre os modelos que têm sido propostos, podem ser mencionados o SQM (*Software Quality Metrics*) [McCa77, Boeh78], o modelo de Gilb [Gilb87] e o GQM (*Goal Question Metric*) [Basi88].

2.3.2 Facetas da Qualidade de Software

McCall *et al.* [McCa77] propuseram um dos primeiros modelos de qualidade de *software*, denominado **SQM** (**Software Quality Metrics – Métricas de Qualidade de Software**), no qual as qualidades almejadas para o produto são estruturadas em uma hierarquia de *fatores*, *critérios* e *métricas*.

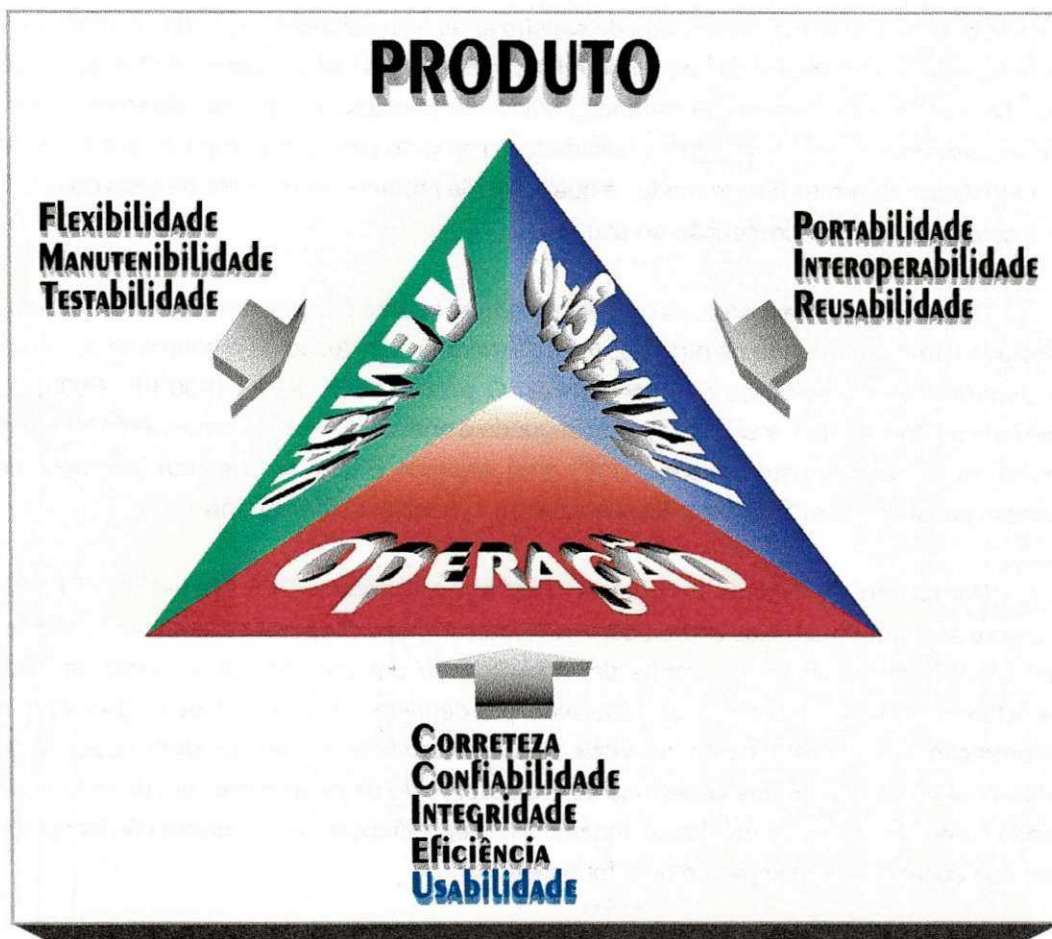


Fig. 2 – Fatores de qualidade de *software* definidos por McCall *et al.* [McCa77].

Os *fatores de qualidade de software* do **SQM** focalizam três aspectos relevantes de um produto de *software* – *características operacionais*, *capacidade de revisão* e *capacidade de adaptação a novos ambientes* (Fig. 2).

Cada *fator de qualidade* representa uma característica comportamental do produto. Cada *critério de qualidade* é um atributo do *fator de qualidade* relacionado com o projeto e com a produção do *software* considerado. Cada *métrica de qualidade* é uma medida associada a algum aspecto de um *critério de qualidade*. Uma ou mais *métricas de qualidade* poderão estar associadas a cada *critério de qualidade*. As setas da Fig. 3 indicam quais dos critérios pré-definidos exercem influência sobre cada um dos *fatores de qualidade* contemplados no modelo.

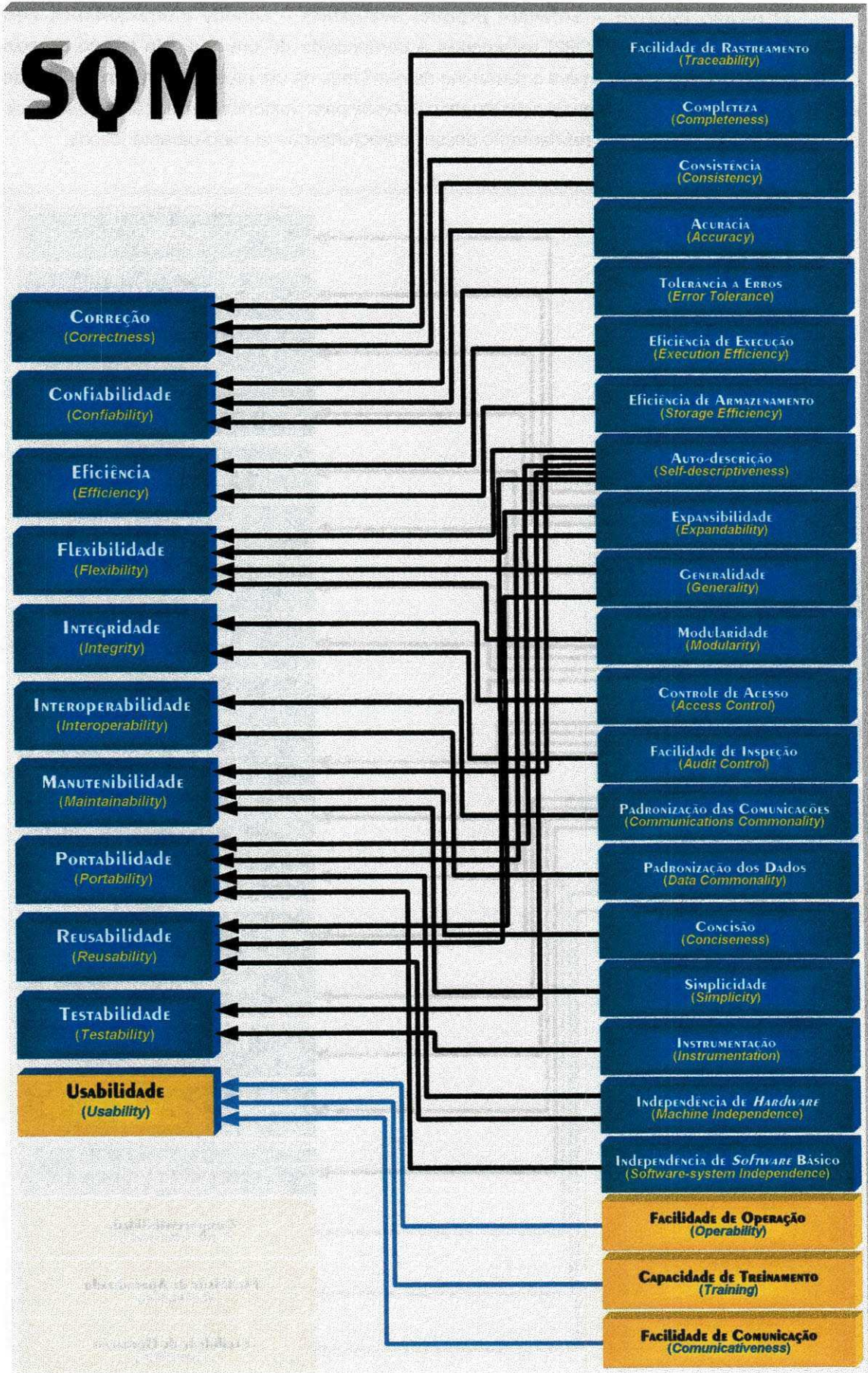


Fig. 3 – Critérios e fatores do modelo SQM [McCa77].

O padrão ISO9126 – **Software product evaluation – Quality characteristics and guidelines for their use** [ISO92] recomenda a composição de um conjunto básico de seis características independentes para a descrição da *qualidade* de um produto, com um mínimo de superposição de atributos, as quais constituem o alicerce para posterior otimização da *qualidade de software*. A Fig. 4 ilustra o refinamento dessas características em sub-características.

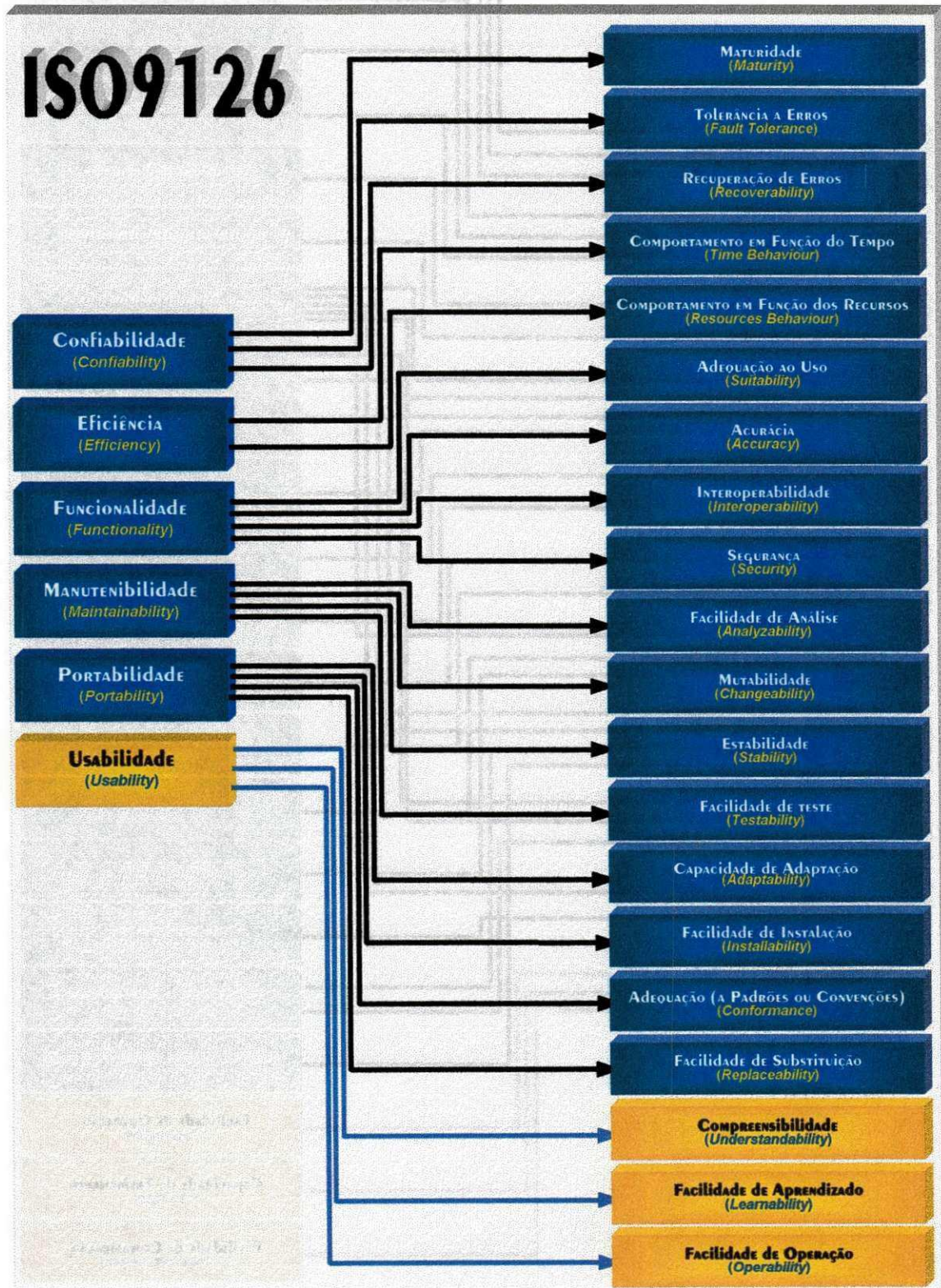


Fig. 4 – Características e sub-características do modelo contido no padrão ISO 9126 [ISO92].

O Quadro 6 apresenta as definições das características do ISO 9126, cujo equivalente brasileiro é a norma da ABNT NBR13596, que aponta igualmente para as mesmas características de *qualidade de software* descritas no Quadro 6.

Quadro 6 – Características de *qualidade de software* do modelo apresentado no documento do padrão ISO 9126 [ISO92].

CARACTERÍSTICA	DEFINIÇÃO
CONFIABILIDADE (<i>Reliability</i>)	Conjunto de atributos relacionados com a capacidade de manutenção do grau de desempenho do produto sob condições pré-estabelecidas, por um período de tempo pré-fixado.
EFICIÊNCIA (<i>Efficiency</i>)	Conjunto de atributos vinculados à relação entre o desempenho do produto e a quantidade de recursos empregados, sob condições pré-estabelecidas.
FUNCIONALIDADE (<i>Functionality</i>)	Conjunto de atributos que implicam a existência de um conjunto de funções e propriedades correspondentes, cada função satisfazendo necessidades pré-estabelecidas ou subentendidas.
MANUTENIBILIDADE (<i>Maintainability</i>)	Conjunto de atributos associados ao esforço necessário para a realização de alterações especificadas, incluindo desde correções, otimizações ou adaptações do produto até alterações de ambiente ou de requisitos e especificações funcionais.
PORTABILIDADE (<i>Portability</i>)	Conjunto de atributos associados à capacidade de migração do produto de um ambiente para outro, considerando-se os contextos <i>organizacional</i> e <i>computacional (hardware e software básico)</i>
USABILIDADE (<i>Usability</i>)	Conjunto de atributos relativos ao esforço associado ao uso do produto e à avaliação de tal uso por um universo pré-estabelecido ou subentendido de usuários.

Vale a pena mencionar que o padrão ISO9126, apesar de recomendar a mensuração direta das referidas características, não sugere métricas nem indica claramente como fazê-lo. Apenas sugere que caso a característica desejada não possa ser mensurada diretamente (especialmente durante a fase de desenvolvimento do produto), deve-se procurar medir algum outro atributo que possa auxiliar o avaliador a prognosticá-la. Além disto, apresenta em um de seus anexos uma indicação de *como* as características podem ser decompostas em sub-características. No entanto, não apresenta diretrizes para a implementação de um bom sistema de prognósticos.

Tervonen [Terv96] descreveu um modelo de *qualidade de software* fundamentado no modelo SQM e no padrão ISO 9126, denominado **Síntese do SQM** (*SQM-synthesis model*). A Fig. 5 mostra uma representação gráfica dos critérios e fatores do modelo concebido por Tervonen [Terv96], ilustrando também a correspondência dos critérios da **Síntese do SQM** ou modelo **GRCM** (*Goal-Rule-Checklist-Metric*) com as características do padrão ISO 9126.



Fig. 5 – Critérios e fatores do modelo Síntese do SQM (ou GRCM) [Terv96] e comparação dos critérios do GRCM com as características do padrão ISO 9126 [ISO92].

Os modelos SQM [McCa77, Boeh78] e ISO9126 [ISO92] diferem basicamente em dois aspectos: (i) a terminologia em geral; e (ii) a estruturação dos "indicadores" de *qualidade*, totalmente hierárquica no padrão ISO (cada sub-característica se relaciona com uma única característica). Além do mais, as sub-características estão relacionadas com aspectos de *qualidade* visíveis ao usuário e não a propriedades internas ao produto analisado.

Modelos como o SQM [McCa77, Boeh78] e o GQM [Basi88, Basi94a] e padrões como o ISO9126 [ISO92] estabelecem uma terminologia para a *qualidade de software*, embora tendam a ser rígidos ou abstratos. Além do mais, não podem assegurar o aprendizado e o uso da terminologia a eles associada, o que pode acarretar a definição insuficiente de estimativas de *qualidade* dentro de um determinado contexto avaliatório, bem como o emprego inadequado e inconsistente de termos de *qualidade* por projetistas e avaliadores [Terv96].

O **GRCM** (*Goal-Rule-Checklist-Metric*), como o autor o denomina, constitui um refinamento do ISO 9126, apresentando *metas (goals)* que se fazem corresponder a *fatores e regras (rules)*, correspondentes a *critérios*. As *listas de inspeção (checklists)* explicam aos avaliadores quais as *regras* a serem discutidas com os projetistas e quais são as *métricas (metrics)* que deverão ser empregadas para implementar a inspeção na prática. A terminologia e o vocabulário das regras deve ser adequado para o desenvolvimento orientado ao objeto. As listas de inspeção são empregadas pelos avaliadores para analisar se o projetista *realmente* seguiu a(s) regra(s).

Dromey [Drom96, Kirc96] desenvolveu uma abordagem que tenta compensar problemas inerentes ao SQM e ao padrão ISO9126, no tocante à carência de base lógica, visando a determinação de *quais* os fatores que deverão ser incluídos na definição de *qualidade* de um produto específico, assim como a identificação de *quais* os critérios específicos que estão relacionados a cada fator selecionado.

2.3.2 Qualidade de Software e Satisfação do Usuário

Os usuários tendem a avaliar a *qualidade* de ferramentas de *software* em termos de sua interação com o produto final [Kirc96]. Um dos estimadores do grau de interação usuário-produto, freqüentemente adotado nos últimos 25 anos, tem sido a *satisfação (subjéitiva) do usuário*, que vem despertando um interesse cada vez maior de pesquisadores de processos interativos usuário-computador. Além do mais, o fato de vir sendo empregado como fator avaliatório do sucesso de processos interativos e do desempenho de sistemas de informação [Meln72] possui uma explicação bastante convincente: o efeito de um bom projeto de interface usuário-computador, que implica estratégias de comunicação bem estruturadas, se traduz tanto no aumento da produtividade do processo interativo [Bail83] quanto na satisfação do usuário e na facilidade de exploração do potencial global do sistema [Neel90].

A satisfação do usuário, por sua vez, associada à idéia de quão apazível seja o uso do sistema [Niel93c], se encontra fortemente atrelada à sua concepção e à avaliação da natureza da aplicação, bem como da interface de comunicação usuário-aplicação; em suma, está intimamente ligada à natureza do produto considerado. Assim, atributos de um produto que contribuem para a satisfação do usuário são um misto de requisitos *funcionais* (ausência ou presença de funções de

seu interesse) e *não-funcionais* (qualidade), assim como das *limitações* do produto (i.e., fatores determinantes de sua aplicabilidade) [Ride91, Kitc96, Kust97]. A distinção entre requisitos funcionais e não-funcionais também é expressa por outros pesquisadores da área, dentre os quais vale a pena mencionar Boehm *et al.* [Boeh78], Cavano e McCall [Cava78], Deutsch e Willis [Deut87], Bemelmans [Beme91] e Delen e Rijsenbrij [Dele92]. Além do mais, o nível de confiança e satisfação do usuário relativo ao uso de um produto depende da atenção dada à maturidade de tal produto, bem como da organização que o desenvolveu, tanto de uma perspectiva externa quanto do ponto de vista dos processos internos de desenvolvimento [Curt93].

Cyert e March ([Cyer63], *in* Bailey [Bail83]) argumentaram que o cotidiano de um usuário de sistemas computacionais impunha uma necessidade de informação que poderia ser atendida ou frustrada conforme as características apresentadas pelo produto em uso. Dez anos depois, Powers e Dickens [Powe73] concluíram, em seu artigo, que o sucesso ou o fracasso de um produto de *software* e, por conseguinte, do uso de um sistema computacional, era ditado sobretudo pela satisfação do usuário. Swanson [Swan74] determinou uma estreita correlação entre a apreciação dos usuários por um produto e a eficiência dos resultados desse produto.

Lucas [Luca75] mostrou uma relação íntima entre o desempenho econômico de uma equipe de vendas e o sistema de informação por esta usado. Evans ([Evan76] *in* [Bail83]) sugeriu a existência de um limite abaixo do qual o usuário evitaria qualquer tipo de interação com o sistema, buscando outras estratégias alternativas de aquisição da informação de interesse. Newman e Segev [Newm80] estudaram o grau de correlação existente entre a reação dos usuários de uma filial bancária a aspectos relativos à satisfação e seu desempenho na organização. Penniman e Dominick [Penn80] resumiram em um quadro de estudos de coleta de informações para avaliação da atitude e comportamento do usuário de sistemas de informação, relacionando estratégias adotadas, trabalhos associados e comentários pertinentes sobre cada estratégia analisada.

Em suma, apesar de nos últimos vinte anos diversos pesquisadores terem investigado a influência da satisfação do usuário na qualidade de produtos interativos, o que se apreende numa leitura revisiva da área é que, na maioria dos casos, não existe um estimador de satisfação que abranja todos os aspectos subjetivos envolvidos pelo contexto, embora se constate em todos eles que o grau de satisfação com relação a um determinado produto apresenta, de algum modo, uma correlação com o sucesso do processo interativo. A visão do usuário, além de relacionar a *qualidade* de um produto à confiabilidade por ele apresentada, vincula a conotação do termo à usabilidade do produto, sobretudo no que diz respeito à *facilidade de instalação, de aprendizado e de uso* [Kitc96].

Numa perspectiva de usabilidade, há duas categorias possíveis de estratégias de estudo da relação *qualidade de software x satisfação do usuário*: (i) *indiretas*, que não estudam propriamente o produto e, por conseguinte, sua interface de usuário, mas apenas as opiniões globais de um universo amostral de usuários sobre o produto em questão; e (ii) *diretas*, que visam avaliar a satisfação do usuário mediante a seleção de atributos adequados ao contexto [Niel93c].

O desenvolvimento de *software* e o projeto de interfaces de usuário, assim como sua supervisão e avaliação, em termos de *qualidade*, se inserem em um domínio mais abrangente, em que projetos e produtos compartilham de características comuns, compondo famílias, atualizadas numa progressão de versões ao longo do tempo. Decisões de projeto tomadas para versões

anteriores provocam efeitos residuais em versões subseqüentes, assim como na concepção de novos produtos derivados. É nesse contexto que se insere uma área cujo *ciclo de vida*²² se estende além do ciclo de vida de qualquer produto, visto que o impacto de suas decisões se projeta na produção futura e nos ciclos de vida dos produtos subseqüentes: a *Engenharia da Usabilidade* [Niel92a].

2.2 Engenharia, Engenharia de Software e Engenharia da Usabilidade

Antes de chegar ao cerne do assunto tratado ao longo deste capítulo, cujos propósitos fundamentais são a discussão da relevância da usabilidade como atributo associado à qualidade de *software*, vale a pena tentar esclarecer um ponto de vista deixado em aberto no capítulo introdutório deste documento – a adequação e a pertinência do uso dos termos *engenharia de software* e *engenharia da usabilidade*.

Sem a pretensão de ser original nem de esgotar a discussão dos termos, que *per si* já constituem temas polêmicos de discussão (às vezes, acirrada) entre os puristas, esta seção deverá ser encarada apenas como um exercício de raciocínio, pertinente ao contexto no qual se insere o tema desta pesquisa, onde se busca apenas esclarecer as conotações em que os termos em questão são empregados neste documento.

2.2.1 Engenharia

Segundo Leite [Leit99], a construção de um artefato pode ocorrer artesanalmente, a partir de processos de tentativa-e-erro, através dos quais se manipula diretamente a matéria-prima com a qual o produto será gerado. Leite [Leit99] acrescenta que os processos de construção artesanal podem evoluir, à medida que incorporam ferramentas específicas e técnicas concebidas e validadas em experiências precedentes.

Tendo em vista que o ato de avaliar é inerente aos seres humanos, conforme discutido no capítulo anterior, a inspeção e identificação de erros e/ou falhas e a verificação da adequação do funcionamento e da utilidade do artefato ocorrerão cedo (durante o desenvolvimento) ou tarde (após o desenvolvimento). A incorporação de ferramentas e técnicas ao desenvolvimento de artefatos facilita o acompanhamento e a análise do processo e, usualmente, conduz a reflexões sobre a atividade anterior: o projeto do artefato. A integração da atividade de projeto ao desenvolvimento de artefatos implica outra etapa de evolução do processo, pois possibilita a visualização do artefato antes mesmo de sua existência enquanto produto, permitindo estender o raio de ação das atividades avaliatórias.

Além disto, os artefatos são desenvolvidos, via de regra, mediante o surgimento de necessidades. Por sua vez, a satisfação de necessidades implica o preenchimento de requisitos. Assim, a incorporação da atividade de projeto ao desenvolvimento de artefatos normalmente aponta para a necessidade de inclusão de outra atividade: a análise de

²² Segundo o padrão 100-1988 do IEEE, o termo refere-se ao período compreendido desde a concepção de um produto até sua retirada do mercado consumidor.

requisitos. Esta atividade deve anteceder o projeto e o desenvolvimento do artefato, a fim de que os objetivos almejados sejam atingidos.

Evidentemente, a evolução do processo de construção de artefatos, traduzida pela incorporação e ordenação de várias atividades, fatalmente atinge um grau de complexidade que suscita a proposição de modelos do processo, através dos quais se pode descrever adequadamente o momento em que cada atividade (análise, especificação, projeto, implementação e avaliação) deverá ser executada. A modelagem do processo estabelece procedimentos metodológicos que descrevem detalhadamente o modo como cada atividade será realizada.

A evolução citada também requer a especialização dos executores de cada atividade pertinente ao seu contexto, exigindo-lhes o conhecimento de um conjunto de princípios e modelos teóricos para o desempenho adequado de suas atribuições. Especialistas em diferentes atividades devem ser agrupados em equipes, lideradas por um gerente, o que exige a análise da estruturação e das estratégias de gestão das atividades relativas ao processo.

Como cada indivíduo representa um papel no modelo de desenvolvimento adotado, também surge a necessidade da divisão coerente das atribuições conferidas a cada um deles ao longo do tempo, ou seja, é necessário o planejamento das tarefas a serem executadas em cada etapa do processo de construção de artefatos. Tendo em vista que todas estas atividades demandarão recursos humanos, físicos, materiais e econômicos, o planejamento deverá incorporar orçamentos de custos, assim como estimativas de prazos para a execução de cada uma das etapas do processo. Afinal de contas, o cliente sem dúvida exigirá uma previsão ou mesmo estipulará um prazo para a entrega do artefato encomendado. Mais uma vez entram em cena os gerentes do processo, a fim de que os recursos sejam bem empregados e os prazos cumpridos.

Quando o desenvolvimento de um artefato incorpora as atividades de análise, projeto, implementação e avaliação, regidas por princípios, modelos, técnicas, ferramentas e métodos específicos e realizadas por equipes de especialistas de acordo com um planejamento que inclui a gestão de orçamentos de custos e estimativas de prazos em todas as etapas do desenvolvimento, pode-se caracterizá-lo como uma *engenharia* [Leit99].

Vale a pena mencionar que o termo *engenharia* deriva-se do latim *ingenere*, que significa "conceber, planejar, criar, construir", por sua vez derivado da palavra *ingenium*, que significa "talento, inteligência", mas também "artefato".

Segundo o *Engineers Council for Professional Development*²³ dos EUA, *engenharia* é a aplicação criativa de princípios científicos (i) no projeto e desenvolvimento de estruturas, máquinas, equipamentos, processos de manufatura ou produtos que os utilize separada ou conjuntamente; ou (ii) na construção ou operação dos mesmos com pleno conhecimento de seu projeto; ou (iii) na previsão de comportamento dos mesmos sob condições específicas de operação, em qualquer um dos casos com uma função planejada, visando economia na operação e segurança para a vida e a propriedade [EB99].

²³ Conselho de Engenheiros para o Desenvolvimento Profissional.

2.2.2 Engenharia de Software

Resta então refletir sobre os processos correntes de desenvolvimento de *software* se enquadram ou não, parcial ou totalmente, nesta definição, o que respaldaria o uso do termo *engenharia de software*. Obviamente se questionará a maturidade de tais processos e, implicitamente, os princípios, métodos e modelos teóricos nos quais se fundamentam. Todavia, é importante não esquecer que vários ramos da engenharia também já estiveram no nível em que se encontra atualmente a área de desenvolvimento de artefatos de *software*, sem contudo deixarem de ser considerados *engenharias*.

Baber [Babe97] afirma que o desenvolvimento de *software* atual se encontra, em muitos aspectos, em uma fase de pré-engenharia, análoga àquela pela qual passaram outros ramos da engenharia tradicional. O autor comenta que na segunda metade do século XIX o campo de tecnologia elétrica, em particular a telegrafia e a telefonia, apresentava problemas análogos àqueles registrados nos processos de desenvolvimento de *software* passados e atuais, acrescentando que falhas consideráveis também já foram lugar comum em muitas outras áreas da engenharia tradicional, e.g., desmoronamento de pontes sob o peso de locomotivas e afundamento de navios por conta da instabilidade dos projetos de seus cascos.

No tocante à telegrafia e telefonia antes e durante a virada do século XIX, Baber [Babe97] comenta que vários fenômenos ainda confundiam os projetistas e limitavam os avanços da Engenharia Elétrica, aos quais se somavam perdas econômicas consideráveis, por conta da compreensão ainda insuficiente de diversos aspectos técnicos relevantes dos sistemas em questão, e.g., a destruição do primeiro cabo transatlântico satisfatoriamente instalado, em 1858; a condução de experimentos onerosos, freqüentemente com resultados inconclusivos, em tentativas para entender fenômenos como o efeito das variações de velocidades de transmissão em circuitos telegráficos bidirecionais e a distorção de sinais de voz em linhas de transmissão.

Baber [Babe97] ainda menciona as iniciativas teóricas de Oliver Heaviside, que solucionou muitos destes problemas através de análises e do uso de lápis e papel, vários anos antes de suas soluções terem sido aceitas e postas em prática pelos "homens da prática", aos olhos dos quais teorias eram consideradas inúteis e, até mesmo, absurdas. Uma delas foi a teoria de Maxwell, que postulava a auto-propagação das ondas eletromagnéticas, questionada como algo que ninguém jamais observara e, evidentemente, jamais observaria, imagine pensar em fazer uso prático dela! Alguns anos mais tarde, Hertz observaria tal fenômeno, fundamentado por considerações teóricas que lhe nortearam no tocante ao que e como investigar. Alguns décadas depois, estas ondas estariam sendo empregadas em comunicações transatlânticas comercialmente viáveis.

A discussão de Baber apenas evidencia a constatação (intuitiva) de que os fatores essenciais responsáveis pela diferença significativa existente entre a precisão dos processos de projeto de ontem e de hoje nos ramos tradicionais da engenharia são (i) a fundamentação científica e matemática sedimentada ao longo de décadas, (ii) a aplicação regular e "natural" desses fundamentos no dia-a-dia e (iii) a capacidade de previsão que

incrementa atualmente a facilidade de descrição dos modelos segundo os quais os artefatos são construídos.

No caso das engenharias civil e mecânica, a fundamentação científica e matemática anteriormente mencionada é fornecida pelas leis de Newton, enquanto que as leis de Maxwell fundamentam a engenharia elétrica). Baber [Babe97] argumenta que uma fundamentação análoga para o desenvolvimento de *software* pode ser encontrada nos estudos de Floyd, Hoare, Dijkstra e outros, embora não tenha ainda um uso amplamente difundido na prática, tendo em vista que esta área do conhecimento ainda vive seus dias de "pré-engenharia", embora usando explicitamente como modelo práticas atuais da engenharia tradicional.

Neste ponto, são acrescentadas algumas definições de autores da área e instituições responsáveis pela padronização de *software*, com o propósito de enriquecer a reflexão sobre o assunto. Bauer [Baue72] define *engenharia de software* como o estabelecimento e uso de princípios consagrados de engenharia para a produção economicamente viável de *software* de qualidade que funcione em máquinas reais.

Na concepção de Fairley [Fair85], *engenharia de software* é um ramo tecnológico e gerencial do conhecimento concernente à produção e manutenção sistemática de produtos de *software* que são progressivamente desenvolvidos e modificados, dentro de estimativas de prazos e custos.

Macro e Buxton [Macr87] referem-se à *engenharia de software* como o estabelecimento e uso de princípios consagrados de engenharia e de práticas de gestão apropriadas, associados à evolução de ferramentas e métodos aplicáveis e ao uso apropriado dessas ferramentas para a obtenção de produtos de *software* de alta qualidade - no sentido mais explícito, dentro de limitações de recursos conhecidas e adequadas.

O *IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries* [IEEE90a] e o padrão IEEE 10.12-1990 [IEEE90b] definem igualmente o termo *engenharia de software* como: (1) a aplicação de uma abordagem sistemática, disciplinada e quantificável no desenvolvimento, operação e manutenção de produtos de *software*, isto é, a aplicação da engenharia ao *software*; e (2) o estudo de abordagens como aquelas definidas em (1).

O *SEI Report on Undergraduate Software Engineering Education* de 1990 [Ford90] diz que enquanto a *engenharia* é a aplicação sistemática do conhecimento científico para a criação e construção de soluções econômicas para problemas práticos da humanidade, a *engenharia de software* é o ramo da engenharia que aplica os princípios da ciência da computação e da matemática para obter soluções economicamente viáveis para problemas de *software*.

Engenharia de software é definida na *Webopedia* [Web00] como o ramo da ciência da computação concernente ao desenvolvimento de aplicações complexas, abrangendo não só os aspectos técnicos de implementação, mas também questões de gestão, tais como a supervisão de equipes de programação, o planejamento e o orçamento de sistemas de *software*.

No tocante aos livros de Engenharia de Software, embora tendam a não apresentarem definições, dedicam seus capítulos introdutórios à explanação de características da área, com comentários que sugerem esboços de definições. Pfleeger [Pfle91] comenta que a *engenharia de software* lida com o projeto e desenvolvimento de produtos de *software* de alta qualidade, sendo responsável pela aplicação de técnicas da ciência da computação na solução de uma ampla variedade de problemas.

Schach [Scha93], por outro lado, afirma que a *engenharia de software* é uma área cujo objetivo é a produção de *software* de qualidade com prazos e orçamentos pré-definidos que satisfaça às necessidades dos usuários.

Segundo Sommerville [Somm97] *engenharia de software* trata da especificação, do desenvolvimento, da gestão e da evolução de sistemas de *software*, elaborando teorias e concebendo métodos e ferramentas aplicáveis ao desenvolvimento de produtos de *software*. O autor acrescenta que embora não lidando com materiais regidos por leis físicas ou processos de manufatura, como outros ramos da engenharia, a *engenharia de software* se insere no contexto da engenharia por lidar com o desenvolvimento de modelos do mundo real.

Gisselquist [Giss98] afirma que desenvolvimento de *software* é engenharia, estabelecendo um rápido, mas não menos interessante, paralelo entre as atividades de gestão, projeto e implementação de produtos nas áreas de desenvolvimento de *software* e de edificações.

Enfim, Mayhew [Mayh99] define *engenharia de software* como um ramo da engenharia relativo ao desenvolvimento de *software* que envolve a definição de requisitos das aplicações, o estabelecimento de metas e atividades de projeto e avaliação em ciclos iterativos até que as metas tenham sido atingidas.

2.2.3 Engenharia da Usabilidade

No contexto mais abrangente dos processos de desenvolvimento de produtos, não apenas de *software*, mas de naturezas as mais diversas, a importância crescente do fator qualidade nas iniciativas de quantificação tanto do processo quanto do produto tem estimulado a emergência de um novo campo de estudo, referido na literatura de língua inglesa como *quality engineering* (engenharia da qualidade). Em nível da qualidade de *software*, a usabilidade tem se tornado uma necessidade competitiva para o sucesso comercial dos produtos de *software*, além de vir sendo incorporado cada vez mais efetivamente nos processos de desenvolvimento, com o propósito de otimizar os produtos através da otimização dos processos através dos quais são desenvolvidos [Butl96a].

Uma década e meia de amadurecimento e sistematização do estudo da usabilidade como atributo decisivo para a aceitação de produtos pelo mercado consumidor resultou no surgimento de uma nova área de investigação, que busca direcionar o processo de desenvolvimento de *software* para a usabilidade dos produtos finais, tanto em nível do processo quanto do produto gerado: a *engenharia da usabilidade*.

Os estudos de Gilb [Gilb84], Shackel [Shac84], Bennett [Benn84, Benn86], Carroll e Rosson [Carr85a] e Butler [Butl85] apresentam uma abordagem de usabilidade unificada, posteriormente denominada *engenharia da usabilidade* por Good *et al.* [Good86], que a definiram como um processo fundamentado na engenharia clássica relativo à especificação, quantitativa e antecipadamente, de *quais* as características e *em que* proporções produtos a serem desenvolvidos deverão apresentar. Segundo Good *et al.* [Good86], sem tal especificação não há outro modo de determinar as necessidades de usabilidade de um produto ou de verificar (qualitativa e quantitativa) se tal produto satisfaz ou não tais necessidades.

A partir desta definição, Good *et al.* [Good86] deixam explícito que a *engenharia da usabilidade* é uma abordagem empregada no desenvolvimento de *software* através da qual a usabilidade de um produto é especificada em termos quantitativos e esta especificação antecede o processo de desenvolvimento do produto. Esta definição também sugere a definição de metas mensuráveis de usabilidade, sem as quais não é possível para o projetista ou o avaliador verificar se o produto final atingiu um grau aceitável de usabilidade (caso atinja as metas pré-fixadas). Hix e Hartson [Hix93] são consonantes com esta definição.

Tyldesley [Tyld88] define *engenharia da usabilidade* como um processo através do qual a usabilidade de um produto é especificada quantitativamente, permitindo que se demonstre se um produto final atinge ou não os níveis exigidos de usabilidade. Este processo inclui medidas objetivas do processo iterativo, definições de modelos de sistemas, do usuário e das interfaces, bem como as interrelações destas entidades, técnicas de projeto cooperativo, técnicas de projeto de interfaces gráficas (GUI), desenvolvimento de diretrizes de projeto, guias de estilo e padrões e atividades de prototipagem [Butl96a].

Butler [Butl96a] comenta que Bennett [Benn84] foi um dos primeiros pesquisadores a empregar o termo *usability engineering*, em meados da década de 80, acrescentando que, após uma década e meia de estudos, a *engenharia da usabilidade* é área técnica que se dedica ao estudo de aspectos referentes ao desenvolvimento de interfaces usuário-computador. Nos, dias atuais, afirma Butler [Butl96a], o processo de desenvolvimento de interfaces é compreendido mais facilmente, apre(e)ndido mais rapidamente e conduzido mais confiavelmente.

Para Lecerof e Paternò [Lece98], *engenharia da usabilidade* é uma área no domínio da interação usuário-computador²⁴ cuja meta fundamental é o desenvolvimento de sistemas usáveis a partir da aplicação, de forma estruturada e sistemática, de diferentes métodos nos diferentes estágios do projeto e do processo de desenvolvimento, assim como da integração de iniciativas de avaliação da usabilidade desde os estágios iniciais do projeto.

Conforme Mayhew [Mayh99], *engenharia da usabilidade* é um processo através do qual são aplicados métodos estruturados no projeto de interfaces usuário-computador visando a usabilidade do produto final.

Como se pôde observar nestes últimos parágrafos, há cerca de quinze anos, *software* "amigável" era, *grosso modo*, qualquer aplicação que interagisse com o usuário através de uma interface de menu ou que fosse dotada de facilidades de recuperação de

²⁴ Human-Computer Interaction (HCI).

erros. Nos dias atuais, a *engenharia da usabilidade* é uma área do conhecimento na qual os pesquisadores e desenvolvedores procuram desenvolver e implementar técnicas que sistematicamente tornem os produtos de *software* mais usáveis, otimizando o produto através da otimização do processo.

Em resumo, além de um exercício de reflexão, esta seção teve o propósito de apresentar diferentes definições encontradas na literatura da área para os termos *engenharia de software* e *engenharia da usabilidade*, além de procurar esclarecer com que conotação estes termos são empregados ao longo de todo o documento. Após ter discutido diferentes conotações do termo usabilidade no capítulo introdutório e de ter mencionado, nas duas últimas seções, a usabilidade como um critério relevante de mensuração da qualidade dos produtos de *software*, torna-se conveniente aprofundar a discussão da importância da usabilidade na aceitação de produtos pelos consumidores, apresentando modelos concebidos nos últimos quinze anos para a mensuração da usabilidade de produtos.

2.3 Usabilidade de Software

Os desenvolvedores de tecnologias bem consolidadas, tais como serviços postais, telefones e televisão alcançaram a meta da usabilidade universal [Shne00]. No entanto, a tecnologia computacional ainda é muito difícil de usar (ou quase inacessível) para muitos indivíduos [Shne98]. Shneiderman [Shne00] cita que em uma pesquisa realizada em um universo amostral de 6.000 usuários de sistemas computacionais verificou-se uma média 5,1 horas por semana despendidas na tentativa de usar computadores, concluindo que se passa mais tempo diante de sistemas computacionais do que ao volante de um automóvel.

Atualmente, as interfaces usuário-computador são entidades muito mais significativas do que costumavam ser há pouco mais de duas décadas, graças à revolução dos computadores pessoais e a redução dos custos do *hardware* que, de um modo geral, vêm tornando os sistemas computacionais mais acessíveis a um número cada vez maior de usuários de áreas de concentração as mais variadas [Hart90, Niel93c, Mand97, Hack98]. Apesar disto, segundo Shneiderman [Shne00], a frustração e a ansiedade dos usuários de sistemas computacionais continua crescendo, enquanto o número de não usuários ainda permanece elevado. O autor acrescenta que embora o barateamento dos custos do *hardware*, do *software* e das redes computacionais prometa fazer muitos novos usuários, a otimização do projeto da interface e do processamento da informação são cruciais para que se atinja graus mais elevados de sucesso.

A diversificação do universo usuário, que implica, sem dúvida, a multiplicidade das tarefas executadas com o auxílio de sistemas de informação automatizados, vem se aliar aos progressos alcançados pela ergonomia e pela engenharia da usabilidade, bem como a oferta de uma diversidade cada vez maior de ferramentas de desenvolvimento e validação de projetos de interfaces. Tais fatores vêm alterando cada vez mais profundamente os rumos da especificação e implementação de interfaces [Laws78] e, por extensão, os alicerces que fundamentam as estratégias avaliatórias destas entidades [Niel89a, Niel89b, Niel90, Kara92, Niel93c, Niel94a, Dix98].

Ao longo das duas últimas décadas, a usabilidade vem sendo cada vez mais considerada como fator [McCa77], característica [ISO92] ou critério [Terv96] de qualidade de *software*. Como se pôde observar na sub-seção 2.1.1, a *usabilidade* figura como uma das componentes de cada modelo de qualidade de *software* apresentado (SQM, ISO9126 e GRM), por sua vez associada a conjuntos de atributos através dos quais pode ser mensurada. Vale a pena revisitar as Figs. 3, 4 e 5 e observar o Quadro 7, que sintetiza os atributos da usabilidade de *software* considerados nos modelos de McCall *et al.* [McCa77], ISO9126 [ISO92] e Tervonen [Terv96].

Quadro 7 – Comparação dos atributos de usabilidade considerados no SQM [McCa77], ISO9126 [ISO92] e GRM [Terv96].

ATRIBUTO DE USABILIDADE	SQM [McCa77]	ISO9126 [ISO92]	GRM [Terv96]
FACILIDADE DE OPERAÇÃO (<i>Operability</i>)	☺	☺	☺
CAPACIDADE DE TREINAMENTO (<i>Training</i>)	☺		
FACILIDADE DE COMUNICAÇÃO (<i>Communicativeness</i>)	☺		
COMPREENSIBILIDADE (<i>Comprehensibility</i>)		☺	
FACILIDADE DE APRENDIZADO (<i>Learnability</i>)		☺	☺
FACILIDADE DE USO (<i>Ease of Use</i>)			☺
APLICABILIDADE (<i>Usefulness</i>)			☺
CONFORMIDADE (<i>Conformance</i>)			☺

Uma análise do Quadro 7 mostra que dos oito atributos de usabilidade considerados nos modelos apresentados na sub-seção 2.1.1 apenas um é compartilhado, a *facilidade de operação* (*operability*). A *facilidade de aprendizado* (*learnability*) aparece em segundo lugar, em termos de atributos compartilhados, já que figura apenas nos modelos do ISO9126 [ISO92] e GRM [Terv96]. Os demais atributos de usabilidade figuram em apenas um dos modelos analisados. Quanto aos pares de atributos (i) *capacidade de treinamento* (*training*) e *facilidade de aprendizado* (*learnability*) e (ii) *facilidade de operação* (*operability*) e *facilidade de uso* (*ease of use*), estes apresentam similaridades entre si, apesar de não serem exatamente equivalentes.

Independentemente de quais atributos tenham sido associados à usabilidade nos modelos de qualidade desenvolvidos até então e da indicação (ou da ausência dela) fornecida por seus conceptores de como mensurar a usabilidade a partir desses atributos, tanto Foley e Van Dam [Fole82] quanto Sulaiman [Sula96] ressaltam que a importância da usabilidade de diversas categorias de sistemas se equipara à funcionalidade deles exigida. Percebe-se, de fato, que nos últimos vinte anos tem-se tentado encorajar os desenvolvedores de *software* a focalizarem a usabilidade de seus produtos tanto quanto se focaliza a funcionalidade ao longo de todo o ciclo de vida de desenvolvimento.

Há 14 anos, Goodwin [Good87] já argumentava que embora usabilidade não fosse um conceito fácil, investir na usabilidade de produtos se afigurava tão importante quanto investir em sua funcionalidade, reforçando que um sistema pouco usável custaria aos usuários, na melhor das hipóteses, o dispêndio de tempo e esforço, enquanto na pior das hipóteses simplesmente não seria usado, visto que a utilidade de suas funções não fora demonstrada. Passada uma década, Sulaiman [Sula96] afirmou que as tentativas de encorajamento haviam tido apenas um sucesso limitado, tendo em vista só haverem atingido algumas companhias, sem alterar fundamentalmente o processo de desenvolvimento de *software* como um todo.







Karat [Kara98] acrescenta que o perfil dos usuários de sistemas de *hardware* e *software* tem mudado, enquanto os sistemas e a cultura de desenvolvimento não tem feito as adaptações necessárias para atender a essas mudanças. Enquanto há uma década apenas especialistas técnicos desenvolviam sistemas para uso próprio ou de outros especialistas técnicos, hoje em dia tais sistemas são ferramentas praticamente onipresentes, desenvolvidas para um segmento crescente de usuários principiantes. Esse contingente de usuários dotados de habilidades e competências em domínios específicos e diversos apresenta pouco ou nenhum interesse em como a tecnologia computacional funciona, desejando apenas ferramentas confiáveis e fáceis de usar, que lhes auxiliem no trabalho diário.

Sobre essas mudanças, Dray [Dray00] chama a atenção para quatro categorias abrangentes de efeitos: (i) efeitos de crescimento da internacionalização da tecnologia; (ii) efeitos da integração crescente da tecnologia; (iii) efeitos demográficos de uso da tecnologia; e (iv) efeitos das mudanças na definição e "posse" do projeto de interfaces de usuário em organizações.

Na verdade, a importância da avaliação da usabilidade de produtos tem sido reconhecida e registrada na literatura da área desde o início dos anos 70 [Mart73, Mill81, Shac90]. A pesquisa sobre a mensuração da usabilidade não só tem contribuído para a elucidação de aspectos relativos à usabilidade [Ride91, Niel94a], como também para a ampliação do grau de conscientização da necessidade de sua integração ao ciclo de vida de desenvolvimento [Kara94b, Sula96, Mand97, Henr98, Mayh99]. Um dos indicadores dessa tomada de consciência é o aumento dos investimentos de muitas companhias de desenvolvimento de *software* em ambientes destinados à avaliação da usabilidade de produtos, equipados com os mais modernos recursos para coleta e análise de dados. Há também a elevação dos investimentos na formação de equipes multidisciplinares de avaliação [Ehr194b, Duma94, Levi96a]. Tais iniciativas também denotam, uma tomada de consciência do fato de que, embora os ensaios de usabilidade não abranjam todos os padrões de uso potencial e funcionalidades do produto analisado, as metas da engenharia da usabilidade dirigem os esforços da equipe de projeto para o usuário e possibilitam a avaliação do trabalho realizado durante o desenvolvimento do produto [Ride91].

O Quadro 8 apresenta alguns aspectos relativos a organizações que implantaram programas de usabilidade no ciclo de desenvolvimento de seus produtos. Trata-se de uma síntese dos estudos de Madsen [Mads99] e Borgholm e Madsen [Borg99], envolvendo seis companhias, três americanas e três dinamarquesas, que adotaram práticas de usabilidade no ciclo de desenvolvimento de seus produtos.

Quadro 8 – Perfís de companhias com programas de usabilidade implantados.

ASPECTO		PERFIS DE COMPANHIAS			PERFIS DE COMPANHIAS		
							
		IBM	AIR	Microsoft	Danfoss	B & O ¹	KMD ²
Produtos desenvolvidos		Software e Hardware computacional	Software para PC, produtos médicos e produtos para clientes	Software para negócios e para clientes, dispositivos de entrada	Válvulas, sensores e controladores eletrônicos para produtos de refrigeração	Sistemas de aquecimento, áudio/vídeo e maquinaria hidráulica	Software para inspeções locais e regionais
Quadro funcional	Fatores Humanos						
	Psicologia						
	Projeto Visual						
	Engenharia						
	Antropologia						
	Ciência da Computação						
Ano de estabelecimento do grupo de fatores humanos		1955	1985	1989	1992	1994	1994
Nº de profissionais na equipe		407	13	117	8	7	9
Nº de empregados da companhia		240.000	400	27.320	18.000	2.288	2.200
Atividades de projeto e avaliação para usabilidade	Estudos de campo						
	Avaliação heurística						
	Prototipagem						
	Testes de projetos						
	Testes de usabilidade						
	Testes pós-lançamento						
Usabilidade no processo de desenvolvimento	Antes (projeto)						
	Após (avaliação)						
	Ao longo						
Canais de comunicação	Relatórios						
	Pôsteres						
	Vídeo						
	Protótipos						
	Encontros						
Participação do usuário	Como testadores						
	Junto aos projetistas						
	Sendo observados						

FONTE: Adaptado de Madsen [Mads99] e Borgholm e Madsen [Borg99].

- 1 Bang & Olufsen
- 2 Kommunedata

LEGENDA

- Elevado, predominante
- Médio, algum
- Baixo, ausente

Em contrapartida, o reconhecimento da importância da engenharia da usabilidade pela maioria das organizações ainda não é bem definido, visto que muitas dessas organizações ainda não conseguem mensurar, em termos financeiros, os benefícios econômicos advindos da adoção de procedimentos e ensaios de usabilidade como atividades diárias. Este quadro ainda é reforçado pelas dificuldades encontradas por muitos profissionais como avaliadores de usabilidade para justificar, com argumentos concretos, os custos conseqüentes de procedimentos que adotam [Kara93, Mayh94]. Acrescente-se a isto o fato de que, mesmo reconhecendo a necessidade e a importância das práticas de projeto centradas no usuário e dos programas de usabilidade adotados pelas grandes e médias corporações de desenvolvimento de *software*, as pequenas organizações ainda encaram tais estratégias como proibitivas, o que lhes impede a inclusão de ensaios de usabilidade no ciclo de desenvolvimento de seus produtos [Rowl94, Brow98].

Neste sentido, diversos pesquisadores (e.g., Aykin [Ayki94], Cox *et al.* [Cox94], Ehrlich e Rohn [Ehr94b], Mayhew [Mayh94, Mayh99], têm demonstrado que os benefícios advindos das estratégias de usabilidade por eles concebidas são compensadores face aos custos associados. Karat [Kara93] descreveu procedimentos de análise de dados de dois projetos, à luz de uma técnica simples de engenharia da usabilidade, que considera o compromisso custos-benefícios. Nielsen [Niel93c] analisou aspectos envolvidos na determinação de custos e benefícios, mostrando que a aplicação da engenharia da usabilidade, em nível empresarial, constitui um bom empreendimento econômico. Rowley [Rowl94] descreveu uma estratégia de planejamento de ensaios de usabilidade no campo (*in situ*), aplicáveis por equipes de desenvolvimento de produtos de *software* a grupos consumidores potenciais separados por grandes distâncias do ambiente de desenvolvimento do produto.

Vários autores (e.g., Nielsen [Niel93b], Beyer e Holtzblatt [Beye94], Brun-Cottan [Brun95], Dumais [Duma95], Stanney *et al.* [Stan97a], Henry [Henr98], Mayhew [Mayh99]) têm sugerido ações que traduzem esforços mais simples de condução de projetos centrados no usuário e, por conseguinte, direcionados para a usabilidade do produto. Dentre as ações mais simples desta natureza, podem ser citadas: (i) diálogos com indivíduos da futura comunidade usuária de um sistema de informação computadorizado; (ii) visitas aos ambientes de trabalho de usuários potenciais e observação/enumeração de quais as atividades que desenvolvem no cotidiano e de como as realizam; (iii) identificação de quais os principais problemas que os usuários potenciais enfrentam e como os solucionam; e (iv) interação mais efetiva com comunidades de usuários potenciais visando a sondagem de soluções mais adequadas de informatização das ferramentas que utilizam em tarefas cotidianas.

Esta é a visão da usabilidade como *abordagem de projeto*. O processo de desenvolvimento voltado para a usabilidade do produto é indispensável em abordagens de projeto centrado no usuário ou projeto participativo, pois insere o usuário no contexto do processo e atenua sua resistência a mudanças tecnológicas que venham a ocorrer em seu ambiente de trabalho, familiarizando-o progressivamente com o produto desde a etapa de análise dos requisitos. A principal limitação das abordagens centradas no usuário reside na escolha de indivíduos representativos da comunidade usuária envolvida no

contexto do projeto, já que não é possível recrutar toda a comunidade. Deste modo, do ponto de vista da maioria dos usuários, o processo de desenvolvimento da interface é desconhecido, não influenciando portanto a formação de sua opinião sobre o produto.

Conforme mencionado no capítulo inicial deste documento, a usabilidade é multifacetada, podendo ser associada a diferentes atributos por indivíduos diferentes, conforme se pôde constatar anteriormente através do estudo dos modelos de qualidade apresentados. A usabilidade também pode ser encarada como *abordagem de projeto*, *atributo de um produto* ou *medida* [Kein99]. Estas três visões apresentam pontos em comum, como ocorre com as regiões de encontro das faces de um poliedro.

Segundo Keinonen [Kein99], muitas organizações vêm passando cada vez mais a encarar a usabilidade por como parte integrante do processo de desenvolvimento de produtos, ao invés de conceber as atividades de avaliação como algo à parte, conduzido por um departamento responsável pela inspeção dos produtos projetados e implementados.

Tal postura tem facilitado a adoção de estratégias de avaliação fáceis de aplicar desde as etapas iniciais do processo de desenvolvimento e capazes de produzir resultados satisfatórios a partir dos recursos disponíveis (e.g., *Discount usability engineering* de Nielsen [Niel95], *Do-it-yourself usability evaluation* de Botman [Botm96], *Quick and dirty* de Thomas [Thom96]). Estas são estratégias que se adequam bem à abordagem integrada de projetos, familiar aos projetistas industriais.

A visão da usabilidade como *atributo de um produto* está associada à exemplificação de propriedades ou características de sistemas ou produtos intimamente relacionadas com sua usabilidade. Tal visão é usualmente explicitada sob a forma de (i) *diretrizes de projeto*²⁵; (ii) *guias de estilo*²⁶; (iii) *padrões*²⁷; e (iv) *listas de princípios de usabilidade*²⁸. Documentos desta natureza são úteis tanto para propósitos de projeto quanto de avaliação, pois apresentam recomendações, definições, diretivas ou normas de propriedades de interfaces em nível genérico ou específico, aplicáveis a diferentes tipos de produtos e contextos de uso. Sua utilização em nível da avaliação de interfaces usuário-computador será explorada mais adiante, no próximo capítulo.

²⁵ E.g., Shneiderman [Shne79, Shne82, Shne84, Shne87], Galitz [Gali85], Smith e Mosier [Smit86b], Dzida [Dzid89], Grudin [Grud89, Grud90, Grud91], Mitchell e Shneiderman [Mitt89], Norcio e Stanley [Norc89], Nickerson e Pew [Nick90], Tognazzini [Togn90], Jeffries *et al.* [Jeff91], Tetzlaff e Schwartz [Tetz91], Karat *et al.* [Kara92], Mayhew [Mayh92], Gillan [Gill93], Gray [Gray93], Queiroz [Quei94], Cohen [Cohe95a], Cohen *et al.* [Cohe95b], Crow [Crow95], Dilli e Hoffman [Dill95], Gorny [Gorn95], Iannella [Iann95], Ogawa e Ueno [Ogaw95], Vanderdonck [Vand95].

²⁶ IBM [IBM91a, IBM91b], Apple [Appl92, Appl97], Microsoft [Micr95], Sun Microsystems [Sun99].

²⁷ ISO [ISO97a, ISO97b, ISO98, ISO99].

²⁸ Nielsen [Niel93b, Niel94a], Johnson [John96b].

Do ponto de vista do usuário, as únicas conseqüências cruciais desses princípios são as características implementadas do produto. Obviamente os produtos adquiridos são avaliados em termos das características apresentadas por suas interfaces, confirmando as opiniões de Hix e Hartson [Hix93] (a *interface* é, para os usuários, "o sistema") e Hackos e Redish [Hack98] (a *interface* é o que os usuários vêem ao usar o produto e o meio através do qual se comunicam com o produto ao usá-lo). Do ponto de vista do projetista e/ou avaliador, as diretrizes de projeto, os guias de estilo, os padrões (corporativos, nacionais ou internacionais) e as heurísticas são instrumentos eminentemente analíticos, que objetivam em primeiro plano servir de respaldo às atividades de projeto ou de inspeção da usabilidade de produtos.

O terceiro modo de focar a usabilidade – como *medida* – diz respeito à mensuração do processo interativo homem-máquina como estratégia destinada à otimização do processo de desenvolvimento ou do produto. Esta é a faceta visualizada na pesquisa ora apresentada, conforme será visto nos próximos capítulos. Sob este ponto de vista, a usabilidade é encarada de um modo que torna os resultados do processo de mensuração o alvo da discussão, o que exige usualmente a adoção de uma abordagem clássica de engenharia da usabilidade ou a proposição de uma nova abordagem para usabilidade.

Shackel [Shac86, Shac91], Nielsen [Niel93b] e a Parte 11 do padrão ISO9241 [ISO98] empregaram abordagens que consideram questões sobre usabilidade em um nível operacional, sobre objetivos de usabilidade e sobre a relação entre a usabilidade, a utilidade, a aceitação de produtos e questões relativas ao processo interativo.

Segundo Shackel [Shac86], a qualidade de um sistema depende de *como* este incorpore e combine vários critérios tradicionalmente vagos e não construtivos no contexto de projeto e implementação. Por sua vez, a *usabilidade* é a "medida" que traduz a concretização de metas especificadas por categorias de usuários interagindo com ambientes específicos, sujeita a vários critérios de conforto, eficiência, eficácia e aceitação. Embora esta definição sugira a quantificação da influência dos termos empregados, tal iniciativa nem sempre constitui uma tarefa trivial ou de resultados facilmente exprimíveis.

2.3.1 Visão de Shackel

A abordagem de Shackel [Shac86, Shac91] se encontra entre as primeiras a reconhecerem o caráter relativo do conceito de usabilidade sob diversos aspectos, tendo sido bastante adotada e modificada (e.g., vide Booth [Booth89], Chapanis [Chap91]). Shackel [Shac86, Shac91] concebeu um modelo de percepção de produtos fundamentado na *aceitação* (*acceptance*), seu conceito de nível mais elevado.

Parte-se do pressuposto que o usuário (consumidor) confronta as propriedades do produto com os esforços necessários para adquiri-lo. Caso decida comprá-lo, a análise de custos-benefícios do processo de aquisição envolverá, de um lado,

a *utilidade* (*utility*), a *usabilidade* (*usability*) e a *capacidade de agradar* (*likeability*) de um conjunto de produtos de interesse e, do outro, os *custos* (*costs*) associados, conforme ilustra a Fig. 6. A melhor alternativa é selecionada, ou seja, aceita, o que torna a *aceitação* dependente da *utilidade*, da *usabilidade*, da *similaridade* e dos *custos* do produto.

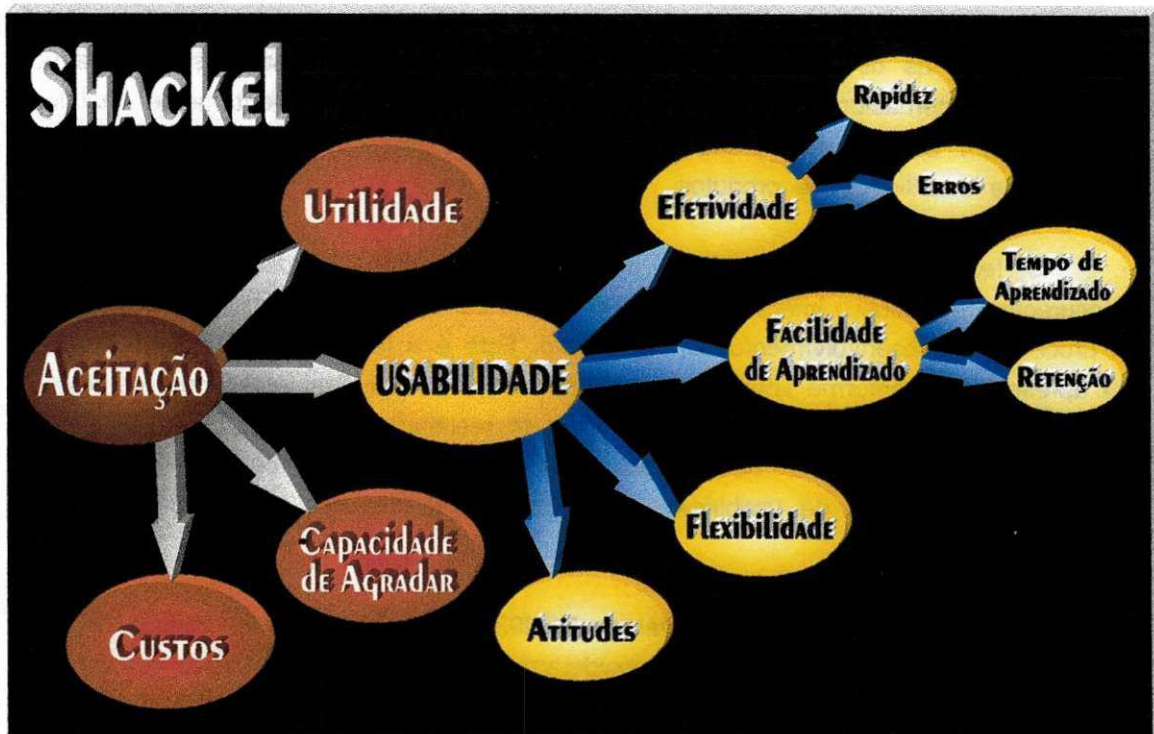


Fig. 6 – Aceitação do produto, conforme Shackle [Shac91], com destaque para as dimensões e medidas concretas da *usabilidade*.

Enquanto a *utilidade* (*utility*) refere-se à confrontação das necessidades do usuário com a funcionalidade do produto, a *usabilidade* (*usability*) diz respeito à facilidade de utilização prática da funcionalidade do produto. Por sua vez, a *capacidade de agradar* (*likeability*) concerne a avaliações subjetivas de cunho emocional/sentimental e os *custos* incluem tanto os encargos econômicos quanto conseqüências sociais e organizacionais do processo de aquisição.

Shackle [Shac91] acrescenta que a *usabilidade* de um sistema ou equipamento é a capacidade de ser usado fácil e efetivamente por um universo especificado de usuários, em termos humanos funcionais, sendo dados treinamento e suporte especificados, de modo a completar um conjunto especificado de tarefas em uma série de cenários ambientais especificados. Destaca-se nesta definição outra mais sucinta - capacidade de ser usado fácil e efetivamente por seres humanos, que se mostra apropriado do ponto de vista da avaliação do produto pelos consumidores, já que o contexto de uso será determinado pela situação específica de cada consumidor. Quanto aos advérbios associados ao modo de uso, *facilmente* refere-se a um determinado nível de análise subjetiva, enquanto *efetivamente* diz respeito a um determinado nível de desempenho humano.

Na concepção de Shackle, usabilidade é uma propriedade não constante de um sistema ou equipamento e sua relatividade é função dos usuários, do treinamento e suporte que lhes foram

dados, das tarefas executadas e dos ambientes onde as tarefas foram concluídas. Dito de outro modo, a usabilidade varia conforme o contexto considerado. Assim, um sistema ou um equipamento será considerado usável se for capaz de atender a uma determinada combinação de usuários, tarefas e ambientes. Deste ponto de vista, a usabilidade apresenta duas facetas: (i) uma associada à percepção subjetiva do produto pelo usuário/consumidor; e (ii) outra relacionada à mensuração objetiva do processo interativo usuário-produto.

Conforme mencionado anteriormente, Shackel apenas sugere os instrumentos, escalas ou aspectos envolvidos no contexto da mensuração objetiva da usabilidade de um produto, não fornecendo indicações sobre *como* conduzi-la. Ciente deste fato, Shackel sugere um conjunto de critérios operacionais a serem satisfeitos em níveis pré-definidos a fim de que um sistema seja considerado *usável*. As escalas sugeridas são apresentadas no Quadro 9.

Quadro 9 – Critérios sugeridos por Shackel [Shac91] para a usabilidade de um sistema.

ESCALA	ESPECIFICAÇÃO
Efetividade (<i>Effectiveness</i>)	Associada aos resultados da interação em termos de rapidez (tempos) e erros.
Facilidade de Aprendizado (<i>Learnability</i>)	Associada à relação entre o desempenho, a quantidade de treinamento e a frequência de uso, i.e., o tempo de aprendizado dos usuários principiantes com um determinado treinamento e retenção do aprendizado por parte de usuários esporádicos.
Flexibilidade (<i>Flexibility</i>)	Associada à possibilidade de adaptação a tarefas e ambientes, além das duas escalas anteriores.
Atitude Mental e Comportamental (<i>Attitude</i>)	Associada a níveis aceitáveis de esforços humanos em termos de fadiga, desconforto, frustração e empenho pessoal.

A concepção de Shackel associa a usabilidade a outros atributos do produto e conceitos de nível mais alto, fornecendo uma definição descritiva do conceito. Direcionada para o contexto complexo de avaliação, sugere critérios de usabilidade concretos e mensuráveis, conforme se pode observar na Fig. 6.

2.3.2 Visão de Nielsen

Nielsen [Niel93b] parte do princípio que, dado um sistema socialmente aceitável, sua *aceitabilidade* (*acceptability*) prática pode ser analisada de diversos ângulos, incluindo aspectos convencionais, e.g. *custo*, *contabilidade*, *compatibilidade com outros sistemas existentes*, além do aspecto *aplicabilidade*²⁹ (*usefulness*), que ele emprega no sentido de o sistema considerado poder ser usado para atingir alguma meta entrevista no contexto em questão. Nielsen, assim como Grudin [Grud92], ainda fraciona a *aplicabilidade* (*usefulness*) em duas sub-categorias de análise: *utilidade*⁶

²⁹ Mesmo em inglês, os termos *usefulness* e *utility* são neologismos neste contexto, cujas definições são pouco distintas, até em bons dicionários da língua inglesa. Daí, a tentativa de traduzi-los como *aplicabilidade* e *utilidade*, palavras existentes no Novo Dicionário Aurélio da Língua Portuguesa [Holl86], que se enquadram no contexto em questão.

(*utility*) e *usabilidade*³⁰ (*usability*). A primeira refere-se ao atendimento de todas as necessidades do usuário pela funcionalidade do sistema, enquanto que a segunda é relativa ao quão aproveitável é tal funcionalidade, i.e., quão bem o usuário pode explorar a funcionalidade do sistema (vide Fig. 7). Este último ponto de vista é consonante com o ponto de vista de Eason [Easo84].

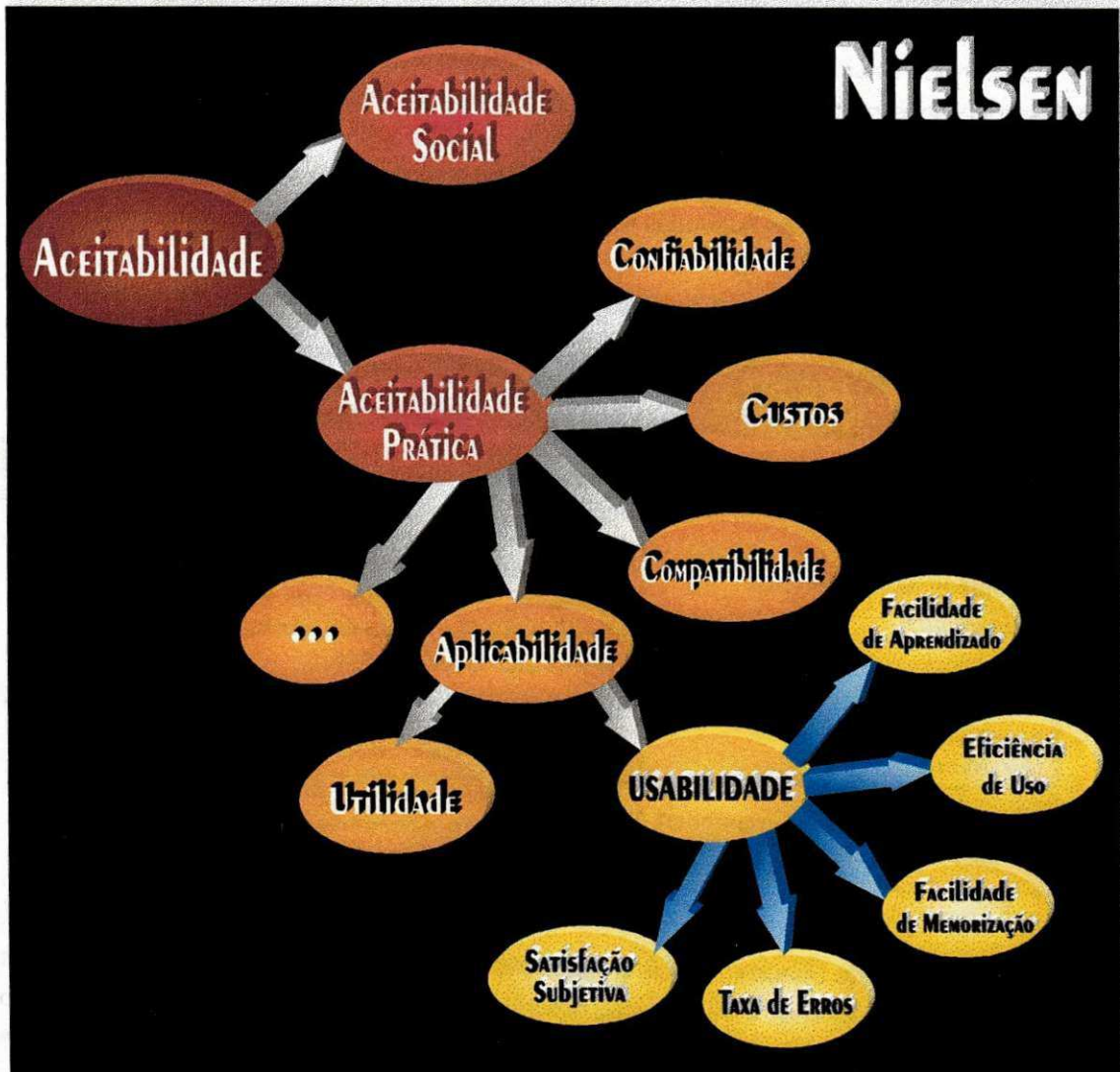


Fig. 7 – Concepção de Nielsen [Niel93c] para a *aceitabilidade* de um sistema, com destaque para as dimensões da *usabilidade*.

Deste modo, verifica-se que o termo *usabilidade* pode ser aplicado, no contexto de Nielsen [Niel90, Niel92a, Niel93b], a quaisquer aspectos de um sistema com os quais usuários interagem, inclusive procedimentos de instalação e manutenção, de onde se pode apreender que tal termo não denota apenas uma propriedade unidimensional da interface de usuário.

³⁰ O termo *usability*, também referido como *useability* ou *usableness*, possui definições genéricas em dicionários correntes da língua inglesa, do tipo "qualidade do que é aproveitável, usável". A tradução deste termo para a língua portuguesa como *usabilidade*, um neologismo técnico, baseou-se em dois fatos: (i) em português, a terminação inglesa *-ility* se faz corresponder, em diversas situações, ao sufixo *-ilidade*; e (ii) da aplicação da regra portuguesa de substantivação de adjetivos terminados em *-ável* (e.g., *miserável* - *miserabilidade*, *afável* - *afabilidade*).

Nielsen não apresenta definições descritivas de usabilidade, embora considere um conjunto de critérios operacionais que emprega para esclarecer adequadamente o conceito. Os critérios operacionais adotados por Nielsen em seu modelo de usabilidade de produtos são listados e especificados no Quadro 10.

Quadro 10 – Critérios da usabilidade de um sistema segundo Nielsen [Niel93b].

CRITÉRIO	ESPECIFICAÇÃO
Facilidade de Aprendizado (<i>Learnability</i>)	Refere-se à capacidade dos principiantes de atingirem rapidamente um nível de desempenho razoável. Considerado por Nielsen como o critério mais básico da usabilidade de produtos, visto que a maioria deles precisa ser fácil de aprender para que possam ser aceitos no mercado, é fundamental nas primeiras experiências de qualquer usuário com sistemas interativos. Alguns sistemas com interfaces <i>difíceis de usar</i> procuram superar tal dificuldade com um bom treinamento.
Eficiência (de Uso) (<i>Efficiency</i>)	Associada ao grau de produtividade do usuário, após haver aprendido a manipular o sistema de informação considerado. Estreitamente relacionada ao desempenho do usuário experiente, se enquadrando em um contexto interpretativo subjetivo e complexo, já que alguns sistemas são tão complexos que exigem de seus usuários vários anos de uso, para que atinjam o desempenho e a habilidade típicos de um usuário experiente. Além do mais, enquanto alguns usuários prosseguem indefinidamente o aprendizado dos sistemas, a maioria normalmente estaciona em um nível de aprendizado que considera "suficiente" para desempenhar as atividades cotidianas [Ross84, Carr87, Carr91].
Memorabilidade (<i>Memorability</i>)	Refere-se à facilidade de memorização dos procedimentos de acesso às facilidades oferecidas por um sistema considerado, evitando todo o processo de re-aprendizado, mesmo quando se trata do uso esporádico do sistema. Tal atributo valoriza não apenas a experiência anterior do contingente de usuários ocasionais - a terceira maior categoria de usuários, além dos principiantes e dos experientes - mas também a experiência daqueles usuários em retorno de qualquer condição temporária que haja interrompido seu contato cotidiano com o sistema com o qual operam.
Taxa de Erros (<i>Errors</i>)	Associada ao cometimento de erros pelo usuário ao utilizar um sistema computacional, assim como aos mecanismos de prevenção contra a ocorrência de erros e à facilidade de correção de tais erros, caso sejam cometidos [Shne87, Shac90, Quel94]. Neste contexto, um erro é definido comumente como qualquer ação que impede o usuário de atingir a meta desejada [Niel93b].
Satisfação Subjetiva (<i>Satisfaction</i>)	Refere-se ao fato de quão agradável é usar o sistema ou, dito de outro modo, quão satisfeito o sistema deixa o usuário ao executar tarefas por este impostas [Bail83, Shne87, Niel93b, Quel94].

2.3.3 Visão da ISO (*International Standards Organization*)

O ISO 9241 é um padrão internacional voltado para o estabelecimento de requisitos ergonômicos para o trabalho empresarial com terminais de visualização. A Parte 11³¹ deste padrão apresenta especificações de usabilidade visando (i) especificações de requisitos de produtos e (ii) avaliação de produtos. Conforme citado no capítulo anterior, a Parte 11 do ISO 9241 [ISO98]

³¹ *Guidance on usability* (Especificações de usabilidade)

define a usabilidade como a abrangência de ação de um produto, segundo a qual este poderá ser empregado por usuários específicos para atingir metas especificadas com *efetividade*³², *eficiência*³³ e *satisfação*³⁴, em um contexto específico de uso. Como *contexto de uso*, a ISO entende usuários, tarefas, equipamentos (*hardware*, *software* e materiais adicionais) e os ambientes físico e social nos quais um produto é usado. A Fig. 8 apresenta uma síntese gráfica dos conceitos relativos à usabilidade tratados no padrão ISO 9241, enquanto a Fig. 9 ilustra as relações entre as dimensões de usabilidade segundo a visão da ISO.



Fig. 8 – Concepção de usabilidade da ISO, através do padrão ISO9241-11 [ISO98].



Fig. 9 – Dimensões da usabilidade de acordo com a Parte 11 do ISO 9241 [ISO98].

³² Precisão e completude com que os usuários atingem metas especificadas.

³³ Recursos despendidos em relação à precisão e completude com que os usuários atingem metas.

³⁴ Ausência de desconforto e exibição de atitudes positivas no tocante ao uso do produto.

Apesar do título (*Usability guidance*), as definições apresentadas nesta parte do padrão ISO 9241 também se aplicam a outras situações que envolvam a interação de usuários com produtos com o propósito de atingir determinados objetivos almejados. Esta abrangência torna a concepção de usabilidade do ISO 9241 um conceito genérico, possibilitando sua aplicação a um leque amplo de situações além de suas aplicações convencionais no âmbito da tecnologia da informação. A Parte 11 teve influência do projeto *MUSIC* do Consórcio Europeu *ESPRIT*, como se constata nos trabalhos de Bevan *et al.* [Beva91, Beva93, Beva94] e Bevan [Beva95a, Beva95b, Beva97].

Bevan e Macleod [Beva94] discutiram uma abordagem de usabilidade fundamentada nas diretrizes do padrão ISO 9241-11, a qual denominaram *qualidade de uso*, definindo-a como uma propriedade do sistema como um todo, o que inclui as práticas de trabalho, a localização e aparência do produto, diferenças individuais entre usuários e outros aspectos que estão intimamente relacionados com o contexto de uso do produto. Segundo este ponto de vista, os atributos de um produto nada mais são do que uma contribuição para a qualidade de uso do sistema como um todo. Por conseguinte, a usabilidade de um produto é encarada como algo dependente não apenas dos usuários e do produto, mas também das tarefas, das metas almejadas, dos equipamentos utilizados e do ambiente de trabalho. Enfim, do contexto de uso.

A Parte 11 do padrão ISO9241 estabelece uma discriminação entre a usabilidade e a qualidade de trabalho a partir da seleção de um ponto de vista específico. Assim, a usabilidade é estudada focalizando-se o produto sob o prisma da qualidade de trabalho. É este aspecto que diferencia fundamentalmente o enfoque do ISO9241 daqueles apresentados por Shackel [Shac86, Shac91] e Nielsen [Niel93b]. Enquanto estes últimos consideram a usabilidade como uma faceta da aceitação do produto pelo consumidor, o padrão ISO 9241 concebe a usabilidade de um produto a partir de um foco especialmente direcionado para a avaliação da qualidade de trabalho. Isto torna os enfoques de Shackel e Nielsen eminentemente centrados no usuário, o que já não ocorre com o enfoque da Parte 11 do padrão ISO 9241.

Como já foi mencionado ao longo desta discussão e também graficamente sintetizado nas Figs. 8 e 9, as dimensões da usabilidade segundo a Parte 11 são: *efetividade*³², *eficiência*³³ e *satisfação*³⁴. Vale salientar que o ISO9241-11 não menciona a experiência prévia do usuário com o produto nem sugere medidas específicas.

Buie [Buie99] comenta que sendo um padrão internacional de processo, as recomendações não poderiam deixar de ser genéricas, o que sugere implicitamente a necessidade de sua adaptação às necessidades de cada organização. Entretanto, o Anexo informativo B do ISO9241-11 apresenta exemplos de medidas de efetividade, eficiência e satisfação tanto quando se focaliza a usabilidade global (*overall usability*) quanto para quando se deseja propriedades específicas para o produto que contribuam para sua usabilidade. No último caso, as propriedades desejadas são associadas a objetivos de usabilidade, que por sua vez são condicionados aos seguintes aspectos: (i) *atendimento das necessidades de usuários treinados*; (ii) *atendimento das necessidades de acompanhamento e uso*; (iii) *atendimento das necessidades de uso intermitente ou infrequente*; (iv) *minimização de solicitações de ajuda*; (v) *facilidade de aprendizado*; (vi) *tolerância a erros*; e (vii) *legibilidade*.

Para cada um destes objetivos de usabilidade, o Anexo B lista vários exemplos de medidas de efetividade, eficiência e satisfação. A experiência do usuário e as medidas de usabilidade são conceitos dissociados no ISO 9241-11. As medidas são categorizadas conforme o aspecto da interação que descrevem, i.e., uma parte das medidas descreve a produção da interação, sua qualidade e quantidade, uma segunda parte mensura os recursos dedicados à interação, sob o ponto de vista da produtividade, enquanto a última parte mensura a percepção subjetiva da interação do ponto de vista do usuário. As medidas têm uma importância secundária do ponto de vista da definição, sendo dada aos avaliadores uma grande liberdade de escolha na seleção das medidas mais apropriadas aos seus propósitos.

Pontos de vista sobre *usabilidade* à parte, pode-se afirmar que *engenharia da usabilidade* é um conjunto de atividades que ocorrem através da vida útil de um produto, como consequência de outra série de atividades significativas realizadas desde antes da concepção da interface usuário-produto [Goul85, Niel93b], o que só vem reforçar a constatação de Grudin [Grud91]: a influência de fatores humanos sobre o sucesso de qualquer produto em desenvolvimento ou mesmo de versões futuras desse produto é um fato.

Thimbleby [Thim94] concorda com a relevância dos critérios de usabilidade da definição de Shackel [Shac86, Shac91] para acordos contratuais de usuários de sistemas com fabricantes, mas questiona sua expressividade junto a atividades desenvolvidas pelos projetistas de sistemas, contrapondo a impossibilidade de uso sistemático de medidas empíricas para a otimização de projetos num contexto desprovido de teorias. Para a expressão *engenharia da usabilidade*, o autor sugere o uso de *usabilidade empírica*, uma abordagem a ser encarada como complementar de outra, por ele adotada, a *usabilidade formal*, embasada em propriedades formais e avaliada por atividades que constituem a *usabilidade de sistematização (formulating usability)*.

No âmbito de tais definições, o encaminhamento de procedimentos avaliatórios de interfaces de usuário "engatilha" processos de descobertas e constatações de inúmeros eventos que ocorrem entre o usuário de teste e o produto sob condições de avaliação, em que o fluxo de informações de interesse parece assumir um ritmo difícil de acompanhar, sobretudo quando não foram traçadas metas sobre os tipos de dados que se deseja coletar.

Não é rara a aquisição de extensos volumes de dados virtualmente sem nenhum valor para o contexto avaliado ou, no melhor dos casos, de contribuição irrisória para a otimização do projeto e da usabilidade da interface considerada. Daí, a necessidade não apenas de planejamento prévio dos procedimentos metodológicos, mas também da previsão de *que* tipo de dados contribuirá mais efetivamente para mensurar e atingir as metas de usabilidade pré-fixadas e *como* deverão ser adquiridos. Eis porque, por extensão, a importância do conhecimento das diferentes técnicas de avaliação consagradas na literatura da área.

Vários métodos formais e informais de avaliação de interfaces usuário-computador têm sido concebidos e experimentados, principalmente nas duas últimas décadas. Alguns deles envolvem usuários em sessões individuais ou grupais, outros são centrados em julgamentos e opiniões de especialistas. Alguns são conduzidos em ambientes onde o alvo da avaliação é usado em condições reais, outros ocorrem em laboratórios, nos quais o alvo da avaliação é testado a partir de tarefas simuladas para refletir contextos reais de uso. Uns focalizam a usabilidade como

princípio de projeto, outros não lidam necessariamente com este atributo. Alguns conservam os procedimentos originais, outros resultam de adaptações a contextos mais específicos.

O próximo capítulo tratará de aspectos relativos aos diferentes métodos de avaliação de processos interativos usuário-computador, destacando suas potencialidades e limitações. O Capítulo 4 complementar a discussão de aspectos relativos à avaliação da usabilidade do Capítulo 3, ressaltando tópicos pertinentes ao contexto dos enfoques avaliatórios consagrados na literatura da área, em especial questões relativas à mensuração da usabilidade e à confrontação de enfoques quantitativos e qualitativos.

3

MÉTODOS DE AVALIAÇÃO DE INTERFACES USUÁRIO-COMPUTADOR

"When the cook tastes the soup, it is formative evaluation;
when the dinner guest tastes the soup, it is summative evaluation."

[Jen Harvey - *Evaluation Cookbook* (1998)]

3	Métodos de Avaliação de Interfaces Usuário-Computador	63
3.1	Métodos de Avaliação de Interfaces Usuário-Computador	65
3.2	Métodos de Avaliação da Usabilidade	68
3.2.1	Ensaio de Usabilidade (<i>Usability Testing</i>)	69
3.2.1.1	Observações (<i>Observation</i>)	72
3.2.1.2	Uso de Questionários (<i>Questionnaires</i>)	75
3.2.1.3	Entrevistas (<i>Interviews</i>)	77
3.2.1.4	Verbalização de Procedimentos (<i>Thinking Aloud</i>)	79
3.2.1.5	Interação Construtiva (<i>Constructive Interaction</i>)	80
3.2.1.6	Ensaio Retrospectivo (<i>Retrospective Testing</i>)	81
3.2.1.7	Captura Automática a partir da Aplicação (<i>Automatic Logging from the Application</i>)	82
3.2.1.8	Discussões em Grupo (<i>Focus Groups</i>)	84
3.2.1.9	Retorno de Opiniões do Usuário (<i>User Feedback</i>) / Ensaio de Usabilidade Remoto (<i>Remote Usability Testing</i>)	86
3.2.2	Inspecões de Usabilidade (<i>Usability Inspections</i>)	91
3.2.2.1	Revisões Sistemáticas (<i>Walkthroughs</i>)	91
3.2.2.2	Diretrizes de Projeto, Guias de Estilo e Padrões (<i>Design Guidelines, Styleguides and Standards</i>)	99
3.2.2.3	Avaliação Heurística (<i>Heuristic Evaluation</i>)	108
3.2.2.4	Inspecção Fundamentada na Perspectiva (<i>Perspective-based Inspection</i>)	114
3.3	Considerações Finais	115

O presente capítulo discute aspectos relativos aos métodos de avaliação de processos interativos usuário-computador, suas potencialidades e limitações. A seção 3.1 – **Métodos de Avaliação de Interfaces Usuário-Computador** – introduz o tema central abordado neste capítulo: as estratégias metodológicas usualmente adotadas na avaliação de interfaces usuário-computador.

A seção 3.2 (**Métodos de Avaliação da Usabilidade**) apresenta uma revisão sobre métodos de avaliação da usabilidade de produtos interativos consagrados pela literatura da área. Esta seção apresenta duas grandes categorias de estratégias avaliatórias da usabilidade de interfaces, a saber: (i) *ensaios de usabilidade*, no contexto dos quais são discutidos nove

métodos de avaliação da usabilidade; e (ii) *inspeções de usabilidade*, estratégias avaliatórias que podem contar com a participação de usuários, embora via de regra envolvam apenas especialistas emitindo julgamentos sobre os produtos avaliados a partir de *revisões sistemáticas, diretrizes de projeto, guias de estilos e padrões e/ou heurísticas de usabilidade*.

Na seção 3.3 (**Considerações Finais**), comenta-se superficialmente sobre os métodos empíricos de avaliação, não fundamentados na usabilidade de produtos interativos, não diretamente relacionados com o escopo deste trabalho. Encerra-se a seção com alguns comentários sobre a aplicação das técnicas revisadas a diferentes contextos avaliatórios, os quais servem de vínculo com as questões pertinentes ao contexto dos enfoques avaliatórios, discutidas no Capítulo 4.

3.1 Métodos de Avaliação de Interfaces Usuário-Computador

De acordo com os capítulos anteriores, as atividades de avaliação de uma interface usuário-computador consistem, em essência: (i) da investigação de problemas ou aspectos questionáveis de projeto e/ou à implementação daquela interface; e (ii) da formulação de soluções consistentes para os problemas encontrados a partir da investigação da interface. Vale a pena lembrar que a interface poderá se encontrar sob a forma de um projeto, um protótipo ou um produto pronto para uso. Assim, os estudos de avaliação poderão envolver esboços, projetos, métodos, aplicações, documentação, treinamento ou ambientes (organizacional, técnico e/ou físico), dependendo da fase em que se encontrar a interface - idéias, concepções de prancheta, simulações computacionais, protótipos de teste, produto a ser documentado, produto para treinamento ou produto final adquirido para compor um contexto de trabalho.

Os estudos avaliatórios podem objetivar a obtenção de resultados que sejam revertidos na otimização da interface investigada. Tais estudos, denominados *formativos (formative)*, são caracterizados fundamentalmente por processos contínuos de investigação, adaptáveis tanto ao progresso global da interface quanto a aspectos específicos, emergentes em decorrência da investigação e da modificação de partes da interface. Em geral, os processos avaliatórios formativos auxiliam a “moldar” o produto final: a interface pronta para uso.

Por outro lado, há também estudos avaliatórios em que os resultados têm o propósito de fornecer diagnósticos globais da interface investigada ao término de diferentes etapas de seu desenvolvimento. Estes diagnósticos se fundamentam usualmente em critérios fixos, podendo fornecer pontos de referência, medidas de sucesso de diferentes etapas do processo de desenvolvimento ou confrontações de metas pré-fixadas e resultados atingidos, ou incluir reações de usuários da interface a investigações comparativas envolvendo outras interfaces similares. Estes estudos são referidos na literatura da área como *somativos (summative)*, como será detalhado adiante.

Independentemente de qual abordagem de questionamento seja adotada, é imprescindível, antes de tudo, determinar se as questões formuladas são as mais adequadas ao contexto da avaliação. Um questionamento inadequado ou pouco realístico implicará, sem dúvida, respostas irrelevantes ou inúteis. Evidentemente, há uma margem de questões que poderão ser e, com certeza, serão alteradas. Por certo outras também poderão ser acrescentadas, a fim de enriquecer o processo de investigação. Entretanto, é crucial haver um eixo de questionamento, a fim de que a avaliação seja produtiva.

Isto exige a concepção de um quadro sinóptico do estudo avaliatório que está sendo proposto, bem como uma justificativa de *como* ele se enquadra no contexto geral do desenvolvimento da interface. É necessário, para tanto um planejamento prévio do processo avaliatório, envolvendo a ponderação de propósitos, interesses, metas a serem atingidas, recursos (físicos, materiais, humanos e financeiros) necessários e disponíveis e limitações práticas. Afinal, há de se convir que cada um destes aspectos apresentará graus de importância diferentes em processos avaliatórios que envolvam, por exemplo, (i) a sondagem preliminar da necessidade de implantação de um determinado serviço ou de desenvolvimento de um dado produto; (ii) a análise do uso de uma versão de teste por uma comunidade de usuários potenciais; e (iii) a

avaliação da interface, sob condições reais de uso, de um produto finalizado.

Outro ponto importante é a consciência da dimensão e do esforço envolvidos, ou seja, se a avaliação exigirá estudos a serem conduzidos em diversas etapas ou se um único estudo satisfará as necessidades do contexto avaliatório. Não menos importante do que o como é o para quem a avaliação será realizada, i.e., se diretamente para um investidor do processo de desenvolvimento ou em nível de um trabalho de pesquisa acadêmica ou de formação de competências na área de avaliação.

A forma de apresentação das informações resultantes do processo avaliatório constitui um aspecto crucial tanto como instrumento de argumentação e justificativa, quanto como meio de persuasão. Diferentes níveis de informação convencem categorias diferentes de indivíduos. A forma como a informação é transmitida também exerce influência na forma como será empregada. Medidas quantitativas, fatos concretos e relatos sucintos podem ser mais úteis para categorias de indivíduos que costumam lidar com indicadores numéricos, e.g., empresários, gerentes de produção, administradores. Por outro lado, informações qualitativas são geralmente mais úteis do que dados numéricos para indivíduos que lidam diretamente com o desenvolvimento do produto, e.g., projetistas, analistas de sistemas, especialistas em fatores humanos.

Embora os recursos disponíveis – infra-estrutura física, equipamentos e recursos humanos (avaliadores e usuários de teste) – exerçam uma influência significativa sobre a quantidade e a qualidade de informação coletada e processada, vale a pena ter em mente que não é apenas um laboratório bem equipado, nem a disponibilidade de recursos financeiros para a contratação de profissionais capacitados e para o recrutamento de usuários de teste que garantirão o sucesso de um processo de avaliação. Obviamente, tais condições poderão atuar no processo como agentes facilitadores. No entanto, a escolha do método mais apropriado ao contexto será ditado pelo propósito da avaliação, o qual, por sua vez, ditará a natureza da investigação, i.e., se o que e como se pretende realizar para que se possa atingir as metas almejadas corresponde a um estudo de natureza *formativa* ou *somativa*, uma vez que a informação a ser coletada exercerá uma influência decisiva nas fontes, ferramentas e técnicas a serem adotadas.

No tocante à metodologia a adotar, o mais recomendável é eleger o método que apresente o raio de abrangência mais extenso e o mínimo de limitações no tocante aos propósitos almejados. Na ausência de um único método que preencha estes requisitos, o mais sensato é integrar dois ou mais métodos complementares, conforme a conveniência, de modo que as respostas fornecidas por um deles substancie aquelas fornecidas pelos demais.

Assim como ocorre com qualquer atividade de investigação, convém coletar o máximo de informações permitidas pelas limitações dos recursos disponíveis, desde que haja um propósito consistente e capacidade para processá-las posteriormente - a fim de atingir os resultados almejados – assim como condições (físicas, materiais, humanas) para refiná-las através de outros estudos avaliatórios, caso a investigação não consiga atender ao questionamento previamente formulado.

A consciência do impacto produzido pelos resultados é outro aspecto a ser considerado no planejamento de um estudo avaliatório, sobretudo em circunstâncias nas quais as respostas obtidas possam gerar pontos de vista conflitantes, que comprometam a otimização ou o sucesso

do produto avaliado. A ética e o sigilo acompanham este aspecto, sendo decisivos para a consolidação do respeito dos demais membros da equipe de avaliação, da segurança dos usuários de teste e da confiança daqueles que solicitaram a investigação. Tais aspectos são fundamentais para o sucesso ou fracasso dos processos de tomada de decisões dependentes dos resultados do estudo avaliatório.

Em suma, os objetivos gerais de qualquer processo avaliatório de interfaces usuário-computador visam: (i) a avaliação das potencialidades do projeto; (ii) a avaliação dos impactos causados pelas decisões de projeto; e (iii) o diagnóstico de problemas relativos ao projeto, independentemente do tipo de interface, do *hardware* e *software* considerados, do estágio do projeto, da presença ou ausência de fatores humanos na avaliação, da abordagem metodológica adotada e do tipo de dados coletados [Lea88, Dix98]. No contexto de uma *abordagem* avaliatória, uma *técnica*³⁵ ou uma combinação de *técnicas* poderá ser adotada, a fim de que o processo seja conduzido segundo os interesses do avaliador [Down91, Niel93b].

Whitefield *et al.* [Whit91] propuseram uma classificação bidimensional de técnicas de avaliação, conforme (i) o *envolvimento ou não de usuários reais* e (ii) a *existência ou não de um resultado de projeto (produto)*. Grosso modo, Dix *et al.* [Dix98] também estabeleceram uma distinção entre a avaliação de um projeto e a avaliação de uma implementação, seja um protótipo ou um produto final. Segundo os autores, no primeiro caso tende-se a focalizar a avaliação pelo projetista, sem o envolvimento direto de usuários. No segundo caso, focaliza-se o uso real do sistema. No entanto, torna-se conveniente atentar para o fato de que esta distinção não é rigorosa, de modo que algumas técnicas podem ser aplicadas em qualquer fase do ciclo de vida do produto, como será visto adiante.

Na verdade, espera-se que os melhores resultados advenham de testes com usuários e sistemas reais, porém nem sempre tais condições são factíveis. Daí, os métodos de prototipagem terem ganho um espaço próprio no contexto avaliatório, em estágios iniciais do projeto, permitindo que os resultados possam, a um só tempo, fornecer as respostas desejadas e, quando necessário, alterar o direcionamento do projeto. Por outro lado, a avaliação heurística também dispensa testes de usuário na análise da usabilidade de um produto.

Da literatura consagrada, apreende-se iniciativas de diversos autores em categorizar técnicas de avaliação e agrupar estratégias avaliatórias afins segundo as categorias acima rotuladas. Tais iniciativas se evidenciam em estudos como aqueles desenvolvidos por Downtown [Down91], Jeffries *et al.* [Jeff91], Miller e Jeffries [Mill92], Nielsen [Niel93b], Nielsen e Mack [Niel94a], Treu [Treu94], Harvey [Harv98] e Dix *et al.* [Dix98], dentre outros.

As seções seguintes descrevem diferentes técnicas de avaliação encontradas na literatura da área, as quais foram agrupadas segundo a categorização geral de Treu [Treu94], por ser mais abrangente. Entretanto, ao longo da apresentação, quando se fizer pertinente, enriquecer-se-á

³⁵ Treu [Treu94] distingue os termos *método* e *técnica*, alegando que uma *técnica* é empregada como um suporte de uma ou mais etapas num procedimento mais geral denominado *método*. Logo em seguida, no entanto, o autor ressalta que ambos são muitas vezes empregados indistintamente. É o que se percebe entre os textos de Jeffries *et al.* [Jeff91] e Miller e Jeffries [Mill92] - onde o termo empregado é *técnica* - e o texto afim de Treu [Treu94] - compilado a partir dos estudos de Jeffries *et al.* [Jeff91] - que utiliza constantemente o termo *método* para descrever a mesma categorização. Neste documento, os termos em questão são empregados indistintamente.

determinados tópicos com abordagens de outros autores, citados à medida que seus pontos de vista forem mencionados.

3.2 Métodos de Avaliação da Usabilidade

Segundo Downtown [Down91], os métodos de avaliação podem ser *analíticos* (técnicas de lápis e papel usadas para aplicar modelos formais de processos interativos a sistemas específicos) ou *empíricos*. Na visão do autor, os métodos avaliatórios empíricos ainda podem ser subdivididos em *informais* (*observação do usuário*) e *formais* (*ensaios elaborados formalmente, questionários e entrevistas*). Como terceira possibilidade, é citado o *julgamento de características específicas de produtos por consultores especialistas em fatores humanos*.

Jeffries *et al.* [Jeff91] e Miller e Jeffries [Mill92] apresentam uma categorização de *métodos de avaliação da usabilidade*, segundo a qual as diversas estratégias definidas, testadas e empregadas correntemente se encontram agrupadas sob quatro rótulos mais abrangentes: *avaliação heurística, uso de diretrizes, revisões sistemáticas cognitivas (cognitive walkthroughs) e ensaios de usabilidade*.

Treu [Treu94] adota na íntegra a categorização de Jeffries *et al.* [Jeff91], no que diz respeito aos *métodos de avaliação da usabilidade*. Entretanto, partindo do fato de que a usabilidade representa apenas um dos princípios de projeto empregados na atualidade, o autor acrescenta novas dimensões ao processo de categorização, reunindo diversas outras estratégias avaliatórias sob dois rótulos mais abrangentes, denominados *métodos de avaliação empíricos e outros métodos*, este último associado a estratégias que fogem ao enquadramento nas duas primeiras categorias mencionadas.

Nielsen [Niel93b] descreve apenas estratégias baseadas na usabilidade do projeto ou do produto, mas dedica um capítulo ao estudo dos *ensaios de usabilidade* (Capítulo 6), bem como trata da *avaliação heurística, de revisões sistemáticas cognitivas (cognitive walkthroughs)* e do *uso de diretrizes* em seções específicas de capítulos (Seções 4.7 e 5.11). Outro capítulo (Capítulo 7) é destinado integralmente ao estudo de um grupo de estratégias por ele denominado *Métodos de Avaliação da Usabilidade além dos Ensaios (Usability Assessment Methods beyond Testing)*. Embora a descrição de Nielsen seja detalhada e abrangente, não se percebe grande preocupação em "rotular" estratégias por categoria, numa acepção mais abrangente do termo.

Cox e Walker [Cox93] descreveram *ensaios de usabilidade*, discutindo diferentes modalidades de *observações*, assim como outros instrumentos de avaliação - *registros automáticos; julgamento por especialistas; uso de testes e questionários; entrevistas; captura automática de indicadores quantificáveis diretamente da aplicação (automatic logging from application); avaliação da estética e uso de padrões e diretrizes de projeto*.

Van Vianen *et al.* [VanV96] enfatizaram a necessidade de esforços combinados dos pesquisadores da área de interação homem-máquina no sentido de padronizar os ensaios avaliatórios de interfaces de usuário, enfatizando o esforço dispendido pela multinacional *Philips* para implantar em todas as suas agências um programa padronizado para a avaliação de interfaces. Os autores apresentaram uma revisão sumariada de técnicas de avaliação de interfaces empregadas em diferentes fases do processo de desenvolvimento de produtos.

Dix *et al.* [Dix98] fazem um nítida distinção entre a avaliação do projeto e a avaliação da implementação (Capítulo 11, seções 11.4 e 11.5), discriminando para o primeiro caso quatro estratégias possíveis - *revisões sistemáticas cognitivas (cognitive walkthroughs)*, *avaliação heurística*, *avaliação baseada em revisões (review-based evaluation)* e *avaliação baseada em modelos*. Para a avaliação da implementação, Dix *et al.* [Dix98] apontam três grupos de métodos, a saber: (i) *experimentais*; (ii) *de observação*, discriminando as técnicas *verbalização de ações (think aloud)*, *avaliação cooperativa* e *análise de protocolos (lápiz e papel, registro de áudio, registro de vídeo, captura automática)*; e (iii) *de sondagem*, discriminando as técnicas *entrevistas* e *uso de questionários*.

3.2.1 Ensaios de Usabilidade (*Usability Testing*)

O método mais fundamental e em, alguns aspectos, insubstituível, consiste em desenvolver ensaios com usuários "reais", visando a aquisição de informações diretamente de seu contexto de trabalho. Esta estratégia possibilita ao avaliador a coleta de fatos relativos a *como* os indivíduos interagem com sistemas computacionais e *quais* dos seus problemas têm relação precisa com a interface "concreta" em questão. *Ensaios de usabilidade* proporcionam tais informações, se direcionados para metas precisas, segundo planos de teste consistentes e bem estruturados. Através de um *ensaio de usabilidade* ou de um conjunto deles, é possível diagnosticar fatores tais como a *eficiência* de tecnologias de desenvolvimento de produtos, as *competências* técnicas e gerenciais dos membros de uma equipe de geração de produtos e o *nível de integração* dos membros de uma equipe de projeto [Duma89, Duma94].

Ensaios de usabilidade consistem basicamente de estudos de um processo interativo usuário-computador específico, em condições "reais" ou "controladas", onde especialistas em interfaces coletam dados sobre eventos relacionados com a interação propriamente dita e problemas afins ocorridos durante o uso da aplicação por uma amostra da comunidade usuária [Jeff91, Mill92, Kara92, Niel93b, Treu94, Dix98]. Envolvem uma gama de técnicas que se diferenciam segundo o grau de especialidade necessário, a formalidade do procedimento e os custos associados [Mack93a, Dix98].

Também merecem atenção especial os aspectos *confiabilidade* e *validade* dos ensaios realizados. O primeiro diz respeito à obtenção de resultados similares após a repetição dos ensaios sob condições similares, enquanto que o último é concernente à relevância dos testes realizados e dos resultados obtidos para a usabilidade do produto quando utilizado em condições "reais", fora de um laboratório (i.e., à veracidade dos resultados obtidos quando confrontados com os aspectos de usabilidade almejados para o projeto em avaliação) [Mand97].

Ambos os aspectos devem ser tratados cautelosamente, visto que a *confiabilidade* dos testes de usabilidade está intimamente vinculada a um sem número de diferenças individuais entre os usuários, ao passo que a *validade* de tais testes mantém uma estreita relação com a seleção adequada dos usuários de teste, assim como com a seleção de tarefas compatíveis com o universo amostral de teste considerado, incluindo nesse contexto limitações de tempo e influências sócio-culturais [Niel93b, Popo95, Dix98].

Outro ponto de reflexão diz respeito às metas do processo avaliatório [Mack93a, Harv98], i.e., se o que se pretende dos ensaios é uma avaliação *formativa* (*formative evaluation*) ou *somativa* (*summative evaluation*) da interface de usuário. Aplicando os conceitos de Scriven [Scri67] ao contexto da interação usuário-computador, uma avaliação *formativa* visa o aprimoramento da interface como parte de um processo iterativo e tem como meta principal a verificação de *quais* aspectos específicos são satisfatórios ou deficientes e de *como* tal conhecimento pode implicar a otimização do projeto da interface [Quei96]. Em contrapartida, uma avaliação *somativa* da mesma interface enfocará uma análise de sua qualidade global, visando uma tomada de decisão ou a emissão de um parecer dentro de um dado ângulo de visão, e.g., decidir pela compra de um dentre um grupo de produtos ou opinar sobre as chances de competitividade deste produto em um mercado que disponibiliza outros produtos com interfaces com aplicações "similares".

O planejamento dos testes - pré-requisito essencial e ponto de partida dos *ensaios de usabilidade* - deve, segundo Nielsen [Niel93b], considerar aspectos relativos: (i) à(s) meta(s) a atingir; (ii) ao local e momento adequados para a aplicação; à duração dos ensaios; (iii) aos recursos computacionais necessários, tanto de *hardware* quanto de *software*, além de detalhes referentes ao estado do sistema no início do teste e a simulação das ações³⁶; (iv) à escolha dos avaliadores; à seleção do universo amostral de usuários de teste; (v) à definição das tarefas a executar; ao estabelecimento de critérios de determinação da conclusão correta das tarefas pré-definidas; (vi) à seleção de mecanismos de consulta e auxílio oferecidos aos usuários de teste; (vii) ao nível de interação entre o usuário de teste e o avaliador, antes, durante e após a realização dos testes; (viii) à natureza dos dados a serem coletados, bem como a natureza da análise a que estes serão submetidos, após a coleta; e (ix) aos critérios adotados para julgamento do produto analisado. O Quadro 11 sumaria as quatro etapas de um ensaio de usabilidade típico, segundo o ponto de vista de Nielsen [Niel93b].

Quadro 11 - Etapas de um *Ensaio de Usabilidade* Típico, segundo Nielsen [Niel93b]

ETAPA	DESCRIÇÃO
Preparação	<ul style="list-style-type: none"> ▪ Verificação de todas as condições de teste (sala, sistema computacional, material necessário em geral, instruções e instrumentos de questionamento, etc.) antes da introdução do usuário de teste no ambiente.
Introdução	<ul style="list-style-type: none"> ▪ Recepção do usuário e explanação breve do(s) propósito(s) do ensaio. ▪ Introdução do(s) procedimento(s) de teste. ▪ Distribuição das instruções e esclarecimento de dúvidas.
Teste	<ul style="list-style-type: none"> ▪ Observação imparcial de todos os eventos interativos usuário-sistema. ▪ Anotação, quando necessária, de detalhes pertinentes ao contexto avaliatório. ▪ Diálogo limitado ao estritamente necessário com apenas um dos avaliadores, conforme acordo pré-fixado.
Questionamento/Análise de Resultados	<ul style="list-style-type: none"> ▪ Apresentação do mecanismo de questionamento (questionário, verbalização de procedimentos ou outra estratégia pré-definida) ao usuário de teste. ▪ Associação de todas as informações coletadas (arquivos armazenados no sistema, anotações, questionários e/ou outro meio) ao usuário correspondente. ▪ Elaboração imediata de um breve relatório do ensaio. ▪ Elaboração de relatório final.

³⁶ Nielsen sugere que as condições de carga e tempos de resposta do sistema/rede não sejam excessivamente lentos ou rápidos, a fim de que sejam simuladas condições realísticas durante a aplicação do teste.

Processos de mensuração e estimação constituem a base de grande parte da pesquisa tradicional no campo dos fatores humanos, sendo também de grande importância para a usabilidade de um produto objetivando avaliar se as metas de usabilidade foram atingidas e na comparação de produtos competidores. O desempenho e a preferência do usuário são quase sempre mensurados/estimados a partir da observação de um grupo de usuários realizando uma série de tarefas pré-definidas em ambiente controlado (ensaio laboratorial) ou desenvolvendo atividades cotidianas similares a tarefas de teste pré-especificadas em ambiente não controlado (ensaio sob condições reais), e da coleta de dados objetivos e subjetivos relativos, respectivamente, a contagens de ações, comportamentos e preferências dos usuários acerca do produto utilizado.

Os dados coletados nesses contextos devem estar fundamentados em planejamentos prévios *do que é significativo mensurar/estimar*, conforme os objetivos do ensaio e o tipo de problema em estudo. Tais dados constituem os indicadores objetivos e subjetivos de qualquer ensaio de usabilidade, também denominados *métricas de desempenho e preferência*.

Dentre as métricas de desempenho ou indicadores objetivos típicos em ensaios de usabilidade são mencionados na literatura [Niel93b, Treu94, Duma94, Rubi94, Dix98, Henr98]:

- ⊙ tempo de execução de tarefa;
- ⊙ número e porcentagem de tarefas completadas corretamente com ou sem assistência;
- ⊙ número e porcentagem de tarefas completadas incorretamente;
- ⊙ razão entre acertos e erros;
- ⊙ tempo de recuperação de erros;
- ⊙ tempo de acesso à informação no manual;
- ⊙ tempo de acesso à informação na ajuda *online*;
- ⊙ número de ações errôneas imediatamente subseqüentes;
- ⊙ número de comandos ou de passos para conclusão de uma tarefa;
- ⊙ número de características do sistema lembradas pelo usuário na fase de *questionamento*;
- ⊙ frequência de uso de mecanismos de ajuda e tempo gasto para consultá-los;
- ⊙ número de acessos aos mecanismos de ajuda que solucionaram problemas do usuários;
- ⊙ proporção de transições do parecer do usuário, de favorável a desfavorável, quanto à aplicação e ao sistema;
- ⊙ número de vezes nas quais o usuário expressa frustração (ou satisfação);
- ⊙ proporção de usuários favoráveis ao uso do sistema/aplicação em teste com relação a outros sistemas/aplicações competidores;
- ⊙ número de situações insolúveis dentro do contexto do teste;
- ⊙ proporção de usuários empregando estratégias de trabalho mais eficientes (no caso de vários modos de execução da tarefa);
- ⊙ quantidade de tempo gasto sem interação usuário-sistema³⁷ e número de vezes em que o usuário se "desvia" da tarefa real.

³⁷ Pode-se distinguir entre os intervalos de tempo que o sistema dispense para completar a ação (*tempo de espera do usuário*) e aqueles em que o sistema aguarda uma ação do usuário (*tempo de espera do sistema*).

Quanto às métricas de preferência ou indicadores subjetivos típicos, são mencionados na literatura [Treu94, Duma94, Rubi94, Henr98]:

- i) Preferências escalonadas e opiniões concernentes à:
 - ⊙ utilidade do produto;
 - ⊙ adequação das funcionalidades do produto às tarefas do usuário;
 - ⊙ atendimento das expectativas do usuário pelo produto;
 - ⊙ facilidade de instalação e customização do produto;
 - ⊙ facilidade de aprendizado do produto;
 - ⊙ facilidade de uso do produto;
 - ⊙ facilidade de realização de tarefas específicas;
 - ⊙ utilidade da documentação *online* e *offline*;
 - ⊙ facilidade de localização da informação de interesse na documentação *online* e *offline*;
 - ⊙ facilidade de compreensão da informação de interesse na documentação *online* e *offline*;
 - ⊙ utilidade dos exemplos contidos nos mecanismos de ajuda;

- ii) Preferência por:
 - ⊙ um protótipo *versus* outro protótipo;
 - ⊙ um produto *versus* produtos concorrentes;
 - ⊙ modelo conceitual corrente do produto *versus* modelo anterior;

- iii) Comentários espontâneos:
 - ⊙ "Como é fácil usar esta opção!";
 - ⊙ "Estou completamente perdido!";
 - ⊙ "Acho que agora tenho que consultar o manual!";
 - ⊙ "Não consigo entender esta mensagem de erro!";
 - ⊙ "Esta opção faz exatamente o que eu quero!"

Os mecanismos de avaliação mais comumente empregados em *ensaios de usabilidade* são:

- ⊙ Observações (*Observation*);
- ⊙ Uso de Questionários (*Questionnaires*);
- ⊙ Entrevistas (*Interviews*);
- ⊙ Verbalização de Procedimentos (*Thinking Aloud*);
- ⊙ Interação Construtiva (*Constructive Interaction*);
- ⊙ Ensaio Retrospectivo (*Retrospective Testing*);
- ⊙ Captura Automática diretamente da aplicação (*Automatic Logging from the application*);
- ⊙ Discussões em Grupo (*Focus Groups*);
- ⊙ Retorno Imediato de Opiniões do Usuário (*User Feedback*).

3.2.1.1 Observações (*Observation*)

A avaliação da usabilidade da interface usuário-computador de um sistema pode implicar descobertas valiosas sobre os usuários do sistema, as formas como estes utilizam as funcionalidades com as quais o sistema foi dotado e aspectos do sistema que podem ser otimizados a fim de tornar o uso mais eficaz e eficiente e os usuários mais satisfeitos.

A forma mais valiosa de retorno de informações sobre o processo de interação usuário-sistema é a observação do usuário utilizando o sistema com algum propósito [Guti00]. A aquisição de informações sobre o contexto de trabalho e o processo interativo usuário-computador de comunidades usuárias implica freqüentemente um conhecimento tácito ou, após um certo tempo, cômico dessas comunidades [Wood96]. Um método de avaliação da usabilidade de interfaces usuário-computador extremamente importante consiste em visitar comunidades usuárias e observá-las trabalhando. O avaliador pode estar situado próximo ou ligeiramente afastado do usuário, porém diante do conjunto usuário-sistema, de modo que o processo de interação possa ser observado em sua totalidade (face e mãos do usuário, dispositivos de entrada e saída de dados, etc.). Nos ensaios típicos, o observador toma notas de detalhes relativos ao desempenho do usuário, podendo empregar um cronômetro ou outro dispositivo de temporização para mensurar aspectos temporais do experimento [Lea88] ou registrar total ou parcialmente a sessão a partir de recursos de áudio e/ou vídeo [Niel93b, Dix98, Niel98, Mayh99].

Cox [Cox93] distingue três facetas do método: (i) *observação cooperativa do usuário*; (ii) *observação em ambiente controlado*; e (iii) *observação em ambiente natural (ou de campo)*. Segundo Cox, a *observação cooperativa do usuário* deverá ser empregada por qualquer projetista durante o desenvolvimento e o teste de um produto. A estratégia consiste na observação, por uma equipe de projetistas e especialistas de usabilidade, de usuários interagindo com um protótipo ou uma versão beta do produto considerado, sendo ou não definidas tarefas específicas. O diálogo usuário-equipe é conduzido por um dos componentes da equipe eleito previamente, a fim de evitar confundir o usuário ou desviar sua atenção do alvo da avaliação.

No processo de *observação em ambiente controlado* sugerido por Cox, que também pode ser denominado *observação laboratorial*, a observação pode ser conduzida em vários níveis, podendo a equipe de observação dividir com o usuário de teste o mesmo espaço do laboratório de usabilidade ou observá-lo através de uma superfície refletora unidirecional. O usuário de teste deve ser instruído sobre as tarefas que deverá desempenhar ao longo do ensaio e verbalizar suas ações enquanto as executa³⁸. Nielsen [Niel93b, Niel97a] sugere, além da tomada de notas, o uso de câmaras de vídeo para registro auxiliar das ações do usuário. As ações do usuário podem fornecer informações as mais diversas, tanto de caráter qualitativo como quantitativo.

Quanto à *observação em ambiente natural*, que também pode ser referida como *observação em condições reais* ou *observação de campo*, Cox afirma que a observação de usuários realizando tarefas em seu ambiente de trabalho, com o auxílio de ferramentas de

³⁸ É conveniente ressaltar que o ato de verbalizar ações e opiniões, sugerido por Cox nessa estratégia, se assemelha bastante ao procedimento adotado em outra estratégia muito difundida no campo da engenharia da usabilidade, denominada neste documento *verbalização de opiniões*, termo aqui empregado correspondentemente ao *thinking-aloud* da literatura inglesa. No entanto, como será visto mais adiante, alguns aspectos de condução dos ensaios de usabilidade, segundo estas estratégias, apresentam diferenças que os individualizam.

software, permitem ao observador compreender alguns dos fatores externos ao laboratório de desenvolvimento que podem causar impacto junto à comunidade usuária, já que é possível verificar se o usuário trabalha sob pressão, se este sofre muitas interrupções durante a execução das tarefas e como as condições físicas do ambiente de trabalho influenciam seu trabalho. Butler e Tahir [Butl96b], assim como Nielsen [Niel98], ressaltam que a observação no contexto de trabalho possibilita a visualização detalhada das ações dos usuários, o que contribui significativamente para a aquisição de dados associados às metas do processo avaliatório. Mayhew [Mayh99] acrescenta que os usuários podem refletir melhor e explicar suas ações enquanto as executam, além do que tais ações rotineiras, executadas sob as condições oferecidas por seu ambiente de trabalho (físicas, materiais, humanas) os estimulam a lembrar de questões que poderiam não vir à tona em outras circunstâncias.

Lea [Lea88] comentou que uma forma menos intrusiva de *observação direta*, embora mais onerosa em termos de aparato material e de qualidade da visualização do processo interativo, consiste em acompanhar todo o processo, oculto por um espelho unidirecional. Segundo o autor, tal artifício permite que mais de um observador visualize o ensaio, de modo a reduzir efeitos de "polarização sistemática" que, porventura, possam ser introduzidos no registro do comportamento do usuário por um único observador. Nos dias atuais, tal estratégia é correntemente adotada pela maioria dos laboratórios de usabilidade do mundo inteiro, conforme registrado e.g. por Nielsen [Niel93b, Niel97a], Dumas e Redish [Duma94] e Queiroz e Turnell [Quei97a].

Sobre a *observação em ambiente natural*, Nielsen [Niel93b, Niel97a] sugeriu a possibilidade de uso de câmaras de vídeo para registro complementar do processo interativo usuário-sistema, embora contraponha a isto o fato de que nem sempre as organizações-alvos permitem o registro em vídeo de suas atividades de trabalho. Além disso, mencionou que a equipe de observação deve esclarecer o papel que irá desempenhar junto à comunidade usuária visitada. Com relação ao registro em vídeo, Nielsen [Niel98] comenta que, em contraste com os estudos de campo, os estudos laboratoriais os superam em nível de controle da situação, tornando possível o estudo de um número maior de aspectos relativos aos testes, dada a maior flexibilidade de experimentação e condução dos registros em vídeo.

Siochi e Ehrich [Sioc91] também já haviam mencionado o uso de vídeo como registro de apoio à análise subsequente das impressões coletadas ao longo de uma sessão de observação, classificando o processo como tedioso e demorado. Downtown [Down91], por outro lado, havia alertado para as desvantagens da geração de enormes quantidades de dados e do dispêndio de esforço e tempo para a condução de análises acuradas.

Ehrlich *et al.* [Ehr94b] descreveram procedimentos de usabilidade adotados pelo grupo de desenvolvimento da *Lotus Development Corp.*, dentre os quais o *uso de vídeo* e o emprego de *observações laboratoriais indiretas* (observadores e usuários sob observação alocados em salas distintas do laboratório de usabilidade).

Diversos outros autores descreveram ensaios de usabilidade, nos quais a *observação direta* foi adotada como técnica de aquisição de dados, e.g. Roberts e Moran [Robe83], Prümper *et al.* [Prüm91], Frese *et al.* [Fres91], Marchionini e Crane [Marc94], Chimera e Shneiderman

[Chim94], McGraw [McGr94], Siochi e Hix [Sioc94] e Doubleday *et al.* [Doub97].

3.2.1.2 Questionários (*Questionnaires*)

A *satisfação subjetiva* pode ser um atributo particularmente importante de usabilidade para sistemas usados fora de ambientes de trabalho, em um âmbito menos formal, e.g. sistemas pessoais, jogos, ficção interativa ou pintura criativa [Virz91, *in* Niel93b], onde o aspecto entretenimento tem um grau de importância superior à rapidez de execução das tarefas, já que a intenção-chave é o prolongamento da diversão [Carr88b]. Todavia, a análise e a mensuração da satisfação subjetiva não é motivada nem importante apenas pelo contexto do entretenimento. Bailey e Pearson [Bail83] afirmaram que a avaliação da satisfação subjetiva do usuário de sistemas computacionais advém (sobretudo) do desejo de aprimorar a produtividade dos serviços computacionais.

Para estimar níveis de estresse e de conforto do usuário, em princípio, pode-se optar por estratégias envolvendo medições psico-fisiológicas objetivas (e.g., EEGS, dilatação de pupila, ritmo cardíaco, condutividade da pele, pressão sanguínea e nível de adrenalina no sangue) ao invés de sondar o seu bem-estar durante a sua interação com o sistema o que, de certo modo, refletiria sua opinião subjetiva sobre o sistema em uso. Entretanto, tais abordagens requerem normalmente o uso de equipamentos ou a coleta de amostras de sangue do usuário, o que implica condições de intimidação e acúmulo de tensão e, por conseguinte, a quebra de uma atmosfera de relaxamento, tão importante para a execução do experimento, invalidando freqüentemente o esforço dispendido para fins de estudos de engenharia de usabilidade [Alle84, Wast90, Schr92, Treu94].

Em virtude de tais circunstâncias, a sondagem da satisfação subjetiva do usuário, a partir da aquisição de informações sobre a sua opinião acerca do sistema, se afigura como uma estratégia alternativa bastante atrativa. Solicitar aos usuários sua opinião sobre algo mediante a *aplicação de questionários* tem sido uma estratégia bastante empregada em diversos âmbitos de interesse. Se do ponto de vista de um único usuário as implicações de tal procedimento podem adquirir uma conotação puramente subjetiva, quando se pondera as opiniões de múltiplos usuários os resultados passam a se enquadrar num contexto objetivo mais amplo [Niel93b].

Por ser uma estratégia qualitativa de coleta de dados referentes a processos interativos, que pode atingir um contingente significativo de usuários [Cox93, Popo95], o uso de questionários possibilita um melhor estudo de diferentes aspectos da usabilidade de um produto. Questionários ganham uma dimensão extra como ferramentas avaliatórias, principalmente quando se trata de estudos que envolvam diretamente a satisfação do usuário e/ou que visem respostas sobre como os usuários interagem com os sistemas e quais as características que particularmente os satisfazem ou desagradam [Niel93b].

Entrevistos na perspectiva da usabilidade, Nielsen [Niel93b] classificou o uso de questionários como uma técnica de avaliação *indireta*, uma vez que não possibilitam o estudo da interface com o usuário diretamente, mas apenas as opiniões de uma comunidade usuária sobre a referida interface. Por outro lado, o autor chama a atenção para o fato desta técnica também poder

ser enquadrada na categoria das técnicas *diretas*, quando visa medir diretamente a satisfação do usuário. Do ponto de vista da condução do processo, Downtown [Down91] sugeriu duas estratégias de aplicação de questionários como instrumentos de avaliação: *questionários conduzidos por um entrevistador*³⁹ (*interviewer-administered questionnaires*) e *questionários auto-dirigidos* (*self-administered questionnaires*).

Os *questionários conduzidos por um entrevistador* (*interviewer-administered questionnaires*) implicam a disponibilidade de entrevistadores treinados e parecem exigir muito mais esforços de condução do que os *questionários auto-dirigidos* (*self-administered questionnaires*), embora apresentem inúmeras vantagens, dentre as quais o melhor controle do processo de aquisição de dados, maior interação com a comunidade usuária e flexibilidade na seleção das questões em função das diferentes categorias de usuários.

Os *questionários auto-dirigidos* (*self-administered questionnaires*) dispendem menos esforço e são mais "diretos" no tocante à condução, já que os próprios usuários se encarregam do preenchimento. Além do mais, podem atingir uma audiência muito mais numerosa [Popo95]. Entretanto, por não contar com assistência no processo interpretativo das questões e também pelo tênue controle do avaliador sobre os respondentes, esta estratégia pode caracterizar-se por um percentual insatisfatório de retorno de exemplares respondidos⁴⁰, um baixo percentual de questões convenientemente respondidas, uma longa duração do processo e maior dificuldade de análise e compreensão dos resultados.

Popowicz [Popo95] atribuiu o baixo percentual de retorno *dos questionários auto-dirigidos* ao interesse despertado em maior grau *apenas* nos usuários muito satisfeitos e muito insatisfeitos. Adicionalmente, o autor atenta para o fato de que muitos usuários poderão escolher proposições "intermediárias" de resposta ou índices "médios" de escalas avaliatórias porventura adotadas, que podem vir a não refletir obrigatoriamente a realidade do contexto avaliado. Downtown ainda alertou que a ambigüidade de encadeamento das questões também pode desvirtuar a qualidade das respostas e, por conseguinte, o processo de avaliação.

Lea [Lea88] comentou que os questionários consistem usualmente de séries curtas de questões sobre um determinado tópico de interesse, de caráter (i) aberto, exigindo do usuário a anotação de sua opinião a respeito do assunto; ou (ii) fechado, permitindo-lhe responder a partir da seleção de uma dentre um conjunto alternativas (formato múltipla escolha) ou indicar o grau de concordância ou discordância sobre um determinado aspecto considerado através de escalas semânticas e/ou numéricas diferenciais.

Vários autores têm apresentado estudos revisivos e/ou elaborado questionários como instrumentos de coleta de dados (e.g. QUIS [Shne87], ASQ [Lewi91a], SUS [Broo96], SUMI [Kira93, Kira94, Kira96], IsoMetrics [Gedi99]).

³⁹ É conveniente atentar para o fato de que tal estratégia apresenta muitas similaridades com o que outros autores denominam de *entrevista*, uma estratégia que será discutida mais adiante.

⁴⁰ Downtown sugeriu que, em média, apenas 40% dos questionários enviados retornam por via postal.

Outros têm descrito resultados de ensaios avaliatórios fundamentados totalmente⁴¹ ou parcialmente⁴² no uso de questionários.

2.2.1.3 Entrevistas (*Interviews*)

Segundo Cox [Cox93], uma *entrevista* é como um questionário interativo, que permite uma exploração mais aprofundada dos aspectos de interesse do avaliador. Esse aspecto se evidencia na descrição de Downtown [Down91] de *questionários conduzidos por um entrevistador*, discutido acima, assim como em experimentos tais como aquele desenvolvido por Chimera e Shneiderman [Chim94]. Cox também afirma que há uma analogia entre esta estratégia avaliativa e o *uso de questionários*, a *captura automática diretamente da aplicação* e as *observações diretas*. Nielsen [Niel93b] reforça a similaridade entre as duas primeiras estratégias, comentando que ambas envolvem o questionamento do usuário e o registro de suas respostas.

Lea [Lea88], assim como Dix *et al.* [Dix98], considerou esta estratégia mais flexível do que um questionário, em termos da sondagem mais aprofundada de aspectos que não são possíveis através do uso de questionários. Além do mais, comentou sobre a influência dessa flexibilidade sobre os diferentes níveis de estruturação das sessões de teste que tal característica lhes confere.

McAteer [McAt98] distingue as *entrevistas* segundo a formalidade em três categorias: (i) *entrevistas abertas padronizadas (standardised open ended interviews)*, que possibilitam a formulação das questões do mesmo modo para uma amostra considerada, mesmo quando envolvem diferentes entrevistadores; (ii) *entrevistas estruturadas ou guiadas (guided ou structured interviews)*, através das quais o entrevistador conduz um questionamento mais formal dos participantes sobre um tema ou temas bem focalizados; e (iii) *entrevistas informais ou coloquiais (informal ou conversational interviews)*, ao longo das quais o entrevistador adapta rapidamente o questionamento, a fim de respeitar diferenças individuais ou acompanhar alterações comportamentais. As *abertas* correm o risco de omitirem informações importantes que não foram antecipadas pelos estruturadores, enquanto as *estruturadas* podem ser cansativas para ambas as partes, por conta do caráter rígido que apresentam. As últimas, por outro lado, podem exigir um dispêndio de tempo excessivo antes de se tornarem uma fonte de informações sistemáticas.

Do ponto de vista da usabilidade, Nielsen [Niel93b] chama atenção para outro aspecto de similaridade entre questionários e entrevistas: ambos constituem métodos *indiretos* de avaliação, visto que não estudam o produto em si, mas apenas opiniões sobre ele, emitidas por uma amostra da comunidade usuária. Entretanto, comenta que quando o aspecto a ser mensurado é a *satisfação subjetiva do usuário*, questionários e entrevistas passam a ser encarados como instrumentos de *avaliação direta*.

⁴¹ E.g. Bailey e Pearson [Bail83], Ives e Olson [Ives84], Shneiderman [Shne87], Chin *et al.* [Chin88], Kirakowski [Kira93, Kira94, Kira96].

⁴² E.g., Dzida *et al.* [Dzid78], Newman e Segev [Newm80], Root e Draper [Root83], Bailey e Pearson [Bail83], Roberts e Moran [Robe83], Ives *et al.* [Ives83], Ives e Olson [Ives84], Shneiderman [Shne87], Chin *et al.* [Chin88], Estevam [Este90], Ebner *et al.* [Ebne90], MacLean ([MacL91], in [Mack93a], [Mcke93]), Frese *et al.* [Fres91], Prümper *et al.* [Prüm91], Jeffries *et al.* [Jeff91], Lundell e Notess [Lund91], Hix [Hix91], Silva [Silv92], Nielsen e Levy [Niel93e], Marchionini e Crane [Marc94], Queiroz [Quei94], Chimera e Shneiderman [Chim94], Kirakowski [Kira93, Kira94, Kira96], Nelson *et al.* [Nels95], Yamada *et al.* [Yama95], Siegel e Bauer [Sieg97], Doubleday *et al.* [Doub97], Hibino e Rundensteiner [Hibi97], Lim e Usma [Lim98], Gomes [Gome99] e Badre e Jacobs [Badr99].

Entrevistas podem ser conduzidas por telefone, implicando menor custo de implementação e maior rapidez de execução e possibilitando: maior frequência nos contatos usuário-avaliador; atenuação do caráter invasivo das visitas de observação direta; e maior especificidade do processo avaliatório [Niel93b, Popo95, McAt98]. Entretanto, as entrevistas conduzidas mais tradicionalmente exigem o deslocamento do usuário para um laboratório de usabilidade ou do avaliador para a instituição onde o usuário trabalha [Niel93b], o que força a equipe de usabilidade, em muitos casos, a restringir o universo amostral a pequenas frações ou reduzir a abrangência do espaço geográfico, a fim de reduzir o dispêndio de tempo e custos [Shne93].

De um modo geral, entrevistas são bastante adequadas para estudos exploratórios para os quais o avaliador ainda não estabeleceu o foco, pois possibilita um re-direcionamento fácil do questionamento feito pelo entrevistador, assim como o aprofundamento de questões surgidas no contexto do processo de sondagem [Dix98]. Além do mais, o contato direto com o usuário resulta na aquisição de sugestões construtivas específicas.

Entrevistas bem estruturadas são excelentes instrumentos auxiliares na definição de processos de trabalho em abordagens de projeto centradas no usuário, permitindo a obtenção de informações primárias (típicas de sondagens iniciais de grupos de interesse) ou secundárias (que refletem o amadurecimento da interação usuário-entrevistador e apresentam uma conotação de esclarecimento de aspectos anteriormente sondados) [McGr94].

Entrevistas (assim como questionários) com questões "em aberto" são de grande valia quando se deseja registrar incidentes ou ocasiões críticas em que o sistema é particularmente deficiente ou surpreendentemente eficaz, já que permitem à equipe de usabilidade adquirir conhecimento sobre as circunstâncias detalhadas de tais incidentes e levar a outros produtos similares a supressão de incidentes enquadrados no "piores caso" e a extensão de benefícios associados a incidentes inseridos no contexto de "melhor caso" (atingindo, desse modo, um número maior de usuários) [Niel93b, Brun95].

Detalhes de *entrevistas* são raramente mencionados em relatos de estudos avaliatórios, o que dificulta reflexões sobre o significado da informação que se pode coletar a partir destas. Como os dados coletados a partir desta estratégia podem ser "desvirtuados" durante o ensaio, visto constituir um processo interativo entre entrevistador e entrevistado (no qual as respostas do entrevistado estão naturalmente sujeitas a "polarizações" introduzidas pelo entrevistador), Lea [Lea88] aconselha o registro de áudio de todo o processo, para posterior transcrição, ao invés de notas escritas no decorrer das sessões de teste.

Diversos ensaios de usabilidade combinam *entrevistas* com outras estratégias avaliatórias, quer em estágios de desenvolvimento e teste de projetos, quer no estágio de uso de produtos finais, e.g. McGraw [McGr94], Marchionini e Crane [Marc94], Carroll e Rosson [Carr95], Siegel e Bauer [Sieg97] e Lim e Usma [Lim98].

3.2.1.4 Verbalização de Procedimentos (*Thinking Aloud*)

A verbalização de procedimentos (*thinking aloud*), também denominada *relatório verbal* (*verbal report*) [Bain79, Eric84], *comentários falados* (*spoken commentaries*) [Down91] e *protocolo verbal* (*verbal protocol*) [Wrig92] pode ser, segundo Nielsen [Niel93b, Niel97a], o método mais valioso e simples da engenharia da usabilidade. O ensaio consiste basicamente em pré-definir uma ou um conjunto de tarefas envolvendo o sistema ou a aplicação, sob condições de teste e solicitar do usuário de teste a verbalização de todos os procedimentos, idéias, encadeamentos lógicos e opiniões indispensáveis à conclusão da tarefa ou do conjunto de tarefas [Lewi82, Niel93b, Niel97a].

Esta tem sido uma estratégia empregada tanto como foco de experimentos (e.g., Berry e Broadbent [Berr90] e Wright e Monk [Wrig91]) quanto como complemento de outras estratégias (e.g., Siochi e Hix [Sioc94]). Originalmente adotada em ensaios psicológicos [Eric84], esta estratégia avaliativa vem sendo cada vez mais usada em ensaios avaliatórios de processos e produtos interativos [Denn90].

Monk *et al.* [Monk93] descreveram uma variante deste método, à qual denominaram *avaliação cooperativa*. Esta estratégia consiste da interação mais efetiva do avaliador com o usuário de teste, que é encorajado a ver o avaliador como um colaborador e não apenas como um elemento neutro no processo. Ao invés de meramente informar ao usuário que este deve descrever os procedimentos realizados para a conclusão das tarefas de teste, o avaliador passa a interagir com esclarecimentos, a formular questões sobre o desenrolar da tarefa executada e a responder a determinados níveis de questionamento do usuário de teste.

A verbalização de encadeamentos lógicos pode se transformar em verbalização de opiniões sobre o alvo da avaliação. O ponto positivo deste fato é a aquisição de informações indiretas sobre a satisfação do usuário através de comentários informais; entretanto, é importante evitar que sejam o foco do ensaio. Os pontos negativos são basicamente dois: (i) o avaliador deve se manter atento para não alterar aspectos da interface avaliada apenas por conta de comentários emitidos por uma minoria de usuários; (ii) os comentários emitidos pelo usuário poderão ser freqüentemente inadequados de um ponto de vista mais abrangente do projeto de interface, o que exigirá do avaliador maior responsabilidade e bom senso na fase de interpretação dos resultados do ensaio.

Dix *et al.* [Dix98] comentam que a principal vantagem da *verbalização de procedimentos* é a simplicidade, o que requer do avaliador um grau de perícia relativamente baixo. Além de proporcionar ao avaliador a compreensão de diversos problemas existentes em uma interface de usuário, este método também pode ser empregado na observação de *como* uma dada aplicação é realmente utilizada. A *verbalização de procedimentos* pode ser também conduzida ao longo de todo o processo de desenvolvimento da interface, respaldada por concepções em papel ou representações simplificadas (*mock-ups*) simuladas visando avaliar aspectos associados aos estágios iniciais de desenvolvimento.

No entanto, Dix *et al.* [Dix98] acrescentam que a informação obtida com a aplicação desta estratégia avaliativa é freqüentemente subjetiva, podendo também ser seletiva, dependendo das tarefas consideradas. Adicionalmente, o processo de observação pode alterar o modo como os

usuários desempenham as tarefas e, por conseguinte, oferecer uma visão polarizada do contexto analisado. O próprio ato de descrever os procedimentos executados divide a atenção do usuário, podendo fazer com que este altere o modo como interage usualmente com o produto através da interface avaliada.

Vora e Helander [Vora95] propuseram uma variante para a verbalização de procedimentos, à qual denominaram *método da instrução (teaching method)*, que consiste, em essência, de um ensaio de usabilidade no qual usuários de teste interagem com um sistema, em sessões iniciais, a partir da realização de uma seqüência de tarefas pré-determinadas, a fim de se familiarizarem com o modo de funcionamento do sistema, o que lhes confere experiência com o alvo do ensaio. Em seguida, o avaliador apresenta a cada usuário de teste "treinado" um usuário de teste "principliante" ao qual o avaliador instrui previamente no sentido de limitar a participação ativa, para não se tornar um "solucionador" ativo de problemas. Instrui-se, então, cada usuário de teste "treinado" a explicar ao "principliante" o funcionamento do sistema, demonstrando-lhe a seqüência de tarefas anteriormente executadas. Enquanto o usuário "treinado" conduz a sessão junto ao "principliante", o avaliador/observador registra os eventos relevantes ao contexto da avaliação. Os autores recomendam a aplicação deste método nas fases de projeto, codificação, teste e uso efetivo.

A literatura da área apresenta várias menções ao uso desta técnica na avaliação de processos interativos homem-máquina, e.g. Berry e Broadbent [Berr90], Wright e Monk [Wrig91], Wright e Converse [Wrig92], Lewis e Mack [Lewi92], Siegel e Bauer [Sieg97] e Gomes [Gome99].

3.2.1.5 Interação Construtiva (*Constructive Interaction*)

A *interação construtiva*, também denominada *aprendizagem a partir do compartilhamento de descobertas (codiscovery learning)* [Kenn89], *investigação do compartilhamento de descobertas, (codiscovery exploration)* [Kemp98] e *investigação com co-participação (co-participation)* [Wils98] é outra variante da *verbalização de procedimentos* que conta com dois usuários operando juntos o produto-alvo da avaliação [O'Mal84, Kenn89, Cox93, Niel93b, Duma94, Rubi94, Niel97a, Kemp98].

A principal vantagem desta estratégia reside no fato de ser uma situação de teste "mais natural" do que a *verbalização de procedimentos* tradicional, além do que durante a verbalização dos encadeamentos lógicos, os usuários de teste interagem para solucionar problemas. Ademais, Hackman e Biers [Hack92] ressaltaram o fato de ser possível coletar muito mais comentários com esta técnica do que com a estratégia-padrão.

Os inconvenientes dizem respeito às diferenças dos usuários de teste no tocante ao grau de aprendizagem e uso de sistemas computacionais, o que pode prejudicar o andamento do ensaio (pontos de vista divergentes podem converter a sessão num vai-e-vem de ações que não conduzem a lugar nenhum) ou até mesmo invalidá-la (incompatibilidade total de opiniões e/ou de gênios podem conduzir a sessão a um impasse laboratorial).

Nielsen [Niel93b, Niel97a] recomenda a *interação construtiva* especialmente para ensaios de usabilidade que envolvam interfaces de aplicações destinadas a crianças, justificando que pode ser mais difícil enquadrá-las num ensaio de verbalização de procedimentos convencional, ou em projetos em que se possa contar com um número maior de usuários de teste, já que esta

estratégia carece do dobro de usuários de teste de um ensaio de verbalização de procedimentos típico. Cox [Cox93] acrescenta que esta estratégia também se aplica a usuários que "emudecem" quando confrontados com um avaliador (a presença de um "colega de experimento" pode estimular-lhe o diálogo).

Strommen [Stro94] desenvolveu um experimento avaliatório do desempenho de crianças usando três interfaces diferentes fundamentadas no uso de *mouse* para controle de tarefas virtuais. O experimento contou com a participação de 94 crianças de ambos os sexos, trabalhando em pares de mesmo sexo em tarefas baseadas numa metáfora de "passeio na floresta" de uma aplicação desenvolvida pela *IBM Corp.* em parceria com o *Children's Television Workshop*. As tarefas consistiam basicamente de um "passeio virtual" pelo cenário de teste e do apontamento/seleção de objetos pré-definidos através de um *mouse*. De acordo com a evolução do experimento, os pares verbalizavam seus procedimentos, opiniões e comentários pertinentes ao contexto da tarefa.

Kemp e van Gelderen [Kemp98] descrevem a metodologia que desenvolveram e aplicam em sessões de *interação construtiva* na *Philips Kommunikations Industrie (PKI, Nuremberg)*, comentando que sua estratégia se assemelha bastante à *aprendizagem a partir do compartilhamento de descobertas (codiscovery learning)* descrita por Kennedy [Kenn89], embora seja adotada na *Philips* para a análise de aspectos utilitários da interface (e.g., mensuração do tempo de conclusão de tarefas, quantificação do número de erros cometidos), enquanto Kennedy focalizou aspectos hedonísticos e cognitivos da interface.

Nodder *et al.* [Nodd99] relatam haverem adotado *interação construtiva* dentre os métodos empregados nas avaliações laboratoriais que conduziram ao longo do desenvolvimento da versão 3 do aplicativo *NetMeeting* da *Microsoft*, seguida de avaliações envolvendo múltiplos participantes, com um enfoque específico nos tipos de usuário para os quais o produto fora inicialmente desenvolvido, i.e., usuários de produtos de *software*, atuantes em ambientes de trabalho baseados em aplicações compartilhadas.

3.2.1.6 Ensaio Retrospectivo (*Retrospective Testing*)

Ensaio retrospectivo é uma estratégia que pressupõe o uso de vídeo em sessões prévias, a partir do qual é possível a coleta de informações adicionais mediante a revisão do registro em vídeo por um usuário de teste [Hewe87, Cons92, Niel93b, Duma94, Rubi94, Niel97a]. Espera-se comentários revisivos mais extensos e detalhados do usuário de teste do que aqueles feitos durante a realização de tarefas inerentes ao ensaio de usabilidade inicial. Além do mais, a fita de vídeo pode ser rebobinada para frente e para trás pelo avaliador ou determinados instantes podem ser "congelados", permitindo-lhe questionar o usuário de teste mais detalhadamente, sem a preocupação de interferência no teste, que já foi essencialmente concluído. Brun-Cottan e Wall [Brun95] denominam tal procedimento *co-visualização*.

Ensaio retrospectivo são valiosos em situações de teste em que se dispõe de um universo amostral reduzido, visto que o avaliador poderá obter mais informações de interesse a partir de questionamentos sobre o experimento inicial, sem ter que repeti-lo exaustivamente. Além disto, o posicionamento estratégico de câmaras de vídeo possibilita o monitoramento visual do

usuário de teste de perspectivas diferentes daquela em que o avaliador se encontra, permitindo o registro de informações extras sobre ações e reações do usuário ao longo da condução do ensaio.

Por outro lado, como despendem muito mais tempo do que o ensaio inicial registrado em vídeo, as sessões de *ensaios retrospectivos* se tornam proibitivas nos casos em que os custos com os usuários são muito altos ou quando os usuários realizam atividades críticas que impossibilitam sua liberação durante todo o período de ocorrência dos ensaios.

Dix *et al.* [Dix98] mencionam uma variante deste método, à qual enquadram na categoria *análise de protocolos* e denominam *registro em vídeo (video recording)*. A estratégia consiste do registro em vídeo da sessão de teste e de análises posteriores das ações do usuário sem a sua presença para o questionamento, que caracteriza o *ensaio retrospectivo* clássico. Para um único usuário de teste por sessão, os autores recomendam o uso de duas câmaras de vídeo, uma para o registro dos eventos ocorridos em tela e outra focalizada na face e mãos do usuário de teste. A primeira poderá ser dispensada se o sistema computacional for instrumentado para a captura automática dos eventos de tela (através de aplicativos como *DRUM* [MacI93, MacI96], *The Observer* [NIT00] e *ErgoLight Lab Tester* [EL00]).

3.2.1.7 Captura Automática a partir da Aplicação (*Automatic Logging from the Application*)

Também denominada *captura automática de dados (data logging)* por Downtown [Down91], *captura automática de uso real do sistema (actual use logging)* por Nielsen [Niel93b], *coleta contínua de dados sobre o desempenho do usuário (continuous user-performance data collection)* por Shneiderman [Shne93], *coleta automática a partir do sistema (system log)* por Gunn [Gunn98] ou *coleta automática a partir do computador (computer logging)* por Dix *et al.* [Dix98], esta estratégia de usabilidade consiste da monitoração e coleta automática de informações estatísticas relativas ao uso do sistema sob avaliação (ou de aplicações nele instaladas).

Segundo Downtown [Down91], os dados monitorados consistem de eventos tais como o acionamento de teclas e ações com o *mouse* que caracterizam entradas do usuário no sistema (implicando respostas correspondentes do sistema), bem como a captura automática em tempo real do relógio do sistema, visando a coleta de informações sobre a temporização dos eventos. Eventos como padrões de uso de um sistema computacional, velocidade de execução de tarefas por usuários, taxas de erros ou frequência de uso de mecanismos de ajuda *online* são relativamente fáceis de monitorar por gerenciadores de sistemas, graças à arquitetura do *software* instalado, que permite incorporar mecanismos de aquisição de dados [Cox93, Shne93]. Tais eventos são extremamente valiosos para a avaliação de novos sistemas adquiridos, de alterações em procedimentos operacionais, de planos de expansão de um sistema, da otimização de um sistema para fins de treinamento e assim por diante [Good85].

Embora a *captura automática de dados* seja usada como uma forma de coletar informações sobre o uso "real" de um sistema após sua aquisição, esta estratégia também pode ser empregada com método suplementar durante ensaios de usabilidade, para a aquisição de dados de interesse mais detalhados [Niel93b]. Capturar automaticamente

informações sobre o uso real do sistema por seus usuários se mostra particularmente útil para fins de usabilidade porque reflete o desempenho dos usuários face as tarefas cotidianas, bem como porque é relativamente fácil coletar dados relativos a um grande número de usuários atuando sob diferentes circunstâncias, de um modo discreto e "transparente" a esses usuários.

Os benefícios advindos da monitoração automática não se estendem apenas aos comandos e outros recursos do sistema de uso corrente. Recursos não utilizados ou raramente acessados poderão ser analisados com propósitos de otimização e/ou de implementação de mecanismos que os tornem mais acessíveis ao usuário [Neal84, Sioc91, Shne93, Niel93b]. Em última instância, as análises feitas sobre tais recursos poderão conduzir à sua substituição por outros mais eficazes ou à sua remoção definitiva.

Um grande inconveniente desta estratégia é que apenas ações diretas podem ser monitoradas, já que nenhuma informação subsidiária é registrada para respaldar a análise de outras atividades de trabalho do contexto avaliado [Down91], i.e., o método pode mostrar o que os usuários fizeram, mas não porque o fizeram [Niel93b]. Eis porque se procura combinar a *captura automática de dados* com outros métodos, tais como entrevistas [Abra77, Sioc91, Niel93b], observações e monitoração por vídeo [Neal84, Sioc91]. Outra dificuldade apresentada refere-se ao volume de dados capturados [Sioc91, Cox93]. Cox recomenda um estudo cauteloso dos propósitos da aplicação do método e das medidas que deverão ser coletadas, antes de adotá-lo.

A *captura automática de dados* é normalmente implementada a partir da instrumentação de componentes de baixo nível do *software* do sistema, tais como *drivers* de teclado e *mouse*, ou modificando partes do *software* aplicativo de interesse [Hamm94]. A segunda estratégia é mais adequada, visto que torna mais fácil a captura de eventos de interesse relativos à aplicação [Niel93b]. A primeira opção pode ser possível em um nível intermediário se o sistema for implementado através de um sistema de gerenciamento da interface de usuário (UIMS - *User Interface Management System*), que controla entradas e saídas enquanto identifica recursos subordinados ao sistema operacional que venham a ser acessados [Olse88].

Uma vez instrumentado o sistema para captura automática, torna-se fácil manter coletas de dados por extensos períodos de tempo. Se facilidades de análise estatística também forem automatizadas, torna-se possível o processo de captura automática de dados como uma estratégia de alerta à equipe de usabilidade sobre quaisquer alterações nas necessidades da comunidade usuária do sistema, traduzidas por mudanças no modo como utilizam o sistema [Niel93b]. Se o sistema instrumentado "roda" em um *mainframe* ou em estações de trabalho com um espaço de arquivos compartilhado, fica fácil coletar dados automaticamente, a intervalos regulares, a partir da simples cópia dos arquivos capturados de cada usuário.

Adicionalmente ao uso estatístico de dados capturados automaticamente, é possível capturar transcrições completas de sessões de usuários, seja para uso em retrospectivas (*playback*) posteriores [Neal84] ou para análises de padrões de uso, tais como a investigação de *quais* os comandos acessados logo após uma situação de erro [Sioc91].

Conforme Gunn [Gunn98], a captura automática de dados é uma medida objetiva que produz dados estatísticos confiáveis relativos a várias questões, tais como padrões de uso, usabilidade de produtos, estratégias de integração e utilidade percebida de produtos que envolvam o uso de sistemas computacionais. Todavia, tal estratégia não abrange o *porquê* das questões, o que implica a necessidade de avaliação adicional das razões que geraram as informações coletadas.

Descrições da adoção desta técnica de avaliação em ensaios de usabilidade foram registradas por Abrams e Treu [Abra77], Neal e Simons [Neal84], Hanson *et al.* [Hans84], Senay e Stabler [Sena87], Siochi [Sioc89], Bradford *et al.* [Brad90], Tetzlaff e Cleveland [Tetz90], Mosteller e Rooney (*in* Chapanis [Chap91]), Siochi e Ehrich [Sioc91], Cohill e Ehrich ([Cohi83], *in* Siochi e Ehrich [Sioc91]), Siochi e Hix [Sioc94], Dieli *et al.* [Diel94], Swallow *et al.* [Swal97], Lecerof e Paternò [Lece98] e Federico [Fede99].

3.2.1.8 Discussões em Grupo (*Focus Groups*)

Esta estratégia avaliatória, de caráter informal, também denominada *group discussions* [Shne93], pode ser empregada para coletar informações sobre as necessidades da comunidade usuária e sobre seus sentimentos e opiniões antes que a interface projetada tenha sido totalmente implementada, e após já haver sido testada durante algum tempo, i.e., entre as etapas de concepção e prototipagem.

Consiste na reunião de seis a nove usuários (McAteer [McAt98] recomenda de seis a doze), numa sessão de cerca de duas horas, com o objetivo de discutir em conjunto novos conceitos relativos ao projeto da interface para uma aplicação de seu interesse, bem como de identificar problemas a ela associados [Niel93b]. Cada grupo é conduzido por um moderador, responsável pela manutenção do foco da discussão sobre problemas de interesse de um modo que se afigure "transparente" para os componentes do grupo. O papel fundamental do moderador é dar às sessões um caráter de "fóruns livres" ou de encontros informais, quando na realidade deve conduzir as discussões segundo um roteiro pré-planejado pela equipe de desenvolvimento abordando problemas específicos que devem ser melhor conhecidos e solucionados.

A partir da lista pré-elaborada de problemas a serem discutidos e das metas pré-definidas para os tipos de informações que deverão ser coletadas, o moderador assume o difícil encargo de manter a discussão no rumo desejado sem inibir o fluxo livre de idéias e comentários dos participantes do grupo. Além do mais, deve assegurar a participação de todos os membros do grupo na discussão, evitando que os pontos de vista de um membro do grupo ganhem destaque perante os pontos de vista do grupo. Após a sessão, a análise dos dados coletados pode ser simplificada se o moderador sintetizar os tópicos-chaves verbalizados pelos participantes do grupo sob a forma de um relatório curto destacando os pontos altos da discussão. Análises mais detalhadas dessa categoria de informações são bem mais difíceis de serem realizadas, além de extremamente demoradas, devido à natureza pouco estruturada das discussões em grupo.

Duas grandes vantagens desta estratégia são a possibilidade de observação de alguns processos de dinâmica de grupo e questões organizacionais, assim como a capacidade de despertar em cada usuário reações espontâneas e idéias construtivas através da interação com os demais participantes do grupo.

Lea [Lea88] interpreta a *discussão em grupo* como uma espécie de entrevista reunindo diversos usuários de teste numa sessão informal e relativamente não estruturada, o que oferece a vantagem de reduzir o tempo de aplicação do teste e, simultaneamente, permite a interação entre os participantes, gerando muito mais comentários e idéias do que ocorre com as entrevistas convencionais.

Van Vianen *et al.* [VanV96] apontam como vantagens desta técnica a descoberta de problemas inesperados, a evidenciação e priorização de aspectos de projeto mais importantes para os usuários e a rapidez na obtenção de informações de um grande número de usuários. Como desvantagens, os autores listam a quantidade reduzida de dados mensuráveis, o embasamento relativamente pouco sólido para tomadas de decisões e a polarização das informações por líderes de opinião.

Por outro lado, McAteer [McAt98] recomenda o uso desta estratégia para a geração de hipóteses, o desenvolvimento de esboços de entrevistas, a identificação de problemas-chaves, o desenvolvimento de temas emergentes (pertinentes ao contexto da discussão), o desenvolvimento de produtos centrado no usuário e a aquisição de informações reflexivas relativas a interpretações parciais de resultados provenientes de etapas do processo de desenvolvimento.

McAteer [McAt98] ainda apresenta duas variantes das *discussões em grupo*: (i) *discussões em grupo retrospectivas (multistage groups)*, nas quais os participantes são requisitados a comparecerem a mais de uma sessão, a fim de que se possa comparar e esclarecer resultados obtidos em sessões anteriores; e (ii) *discussões em grupo mistas (second-order groups)*, realizadas com uma mescla de participantes de diferentes sessões anteriormente realizadas, nas quais a flexibilidade da estruturação é maior, visando o nivelamento de temas anteriormente discutidos com cada grupo considerado.

Salvador e Howells [Salv98] apresentaram no CHI'98 uma variante da discussão em grupo, denominada *focus troupe*, através da qual são apresentadas vinhetas dramáticas a uma audiência de clientes potenciais. Nessas vinhetas, o conceito de novo produto é caracterizado apenas como um elemento dramático, mas não como um produto tecnológico existente. A vinheta apresenta situações familiares ou típicas, onde os pormenores diferem em função da nova concepção, destacando contextualmente o conceito novo contra um fundo familiar e comum. Segundo os autores, este método é mais eficiente e rápido de aplicar do que a discussão em grupo tradicional, pois a apresentação atrativa, o fundo comum e nenhuma necessidade de conceitos concretos para o produto, oferecem recursos para a formulação de comentários relevantes de clientes em potencial sobre produtos que ainda não existem.

Sato e Salvador [Sato99] descreveram detalhadamente a metodologia da *focus troupe*, cujas sessões se assemelham estruturalmente àquelas de discussões em grupo tradicionais, durando cerca de 2 horas e reunindo em torno de 20 indivíduos, divididos em grupos de 4 a 5 membros. Todavia, o que difere fundamentalmente entre as duas estratégias é que, ao invés de conduzir uma apresentação do produto, como ocorre usualmente nas discussões em grupo, o moderador da *focus troupe* introduz o contexto no qual se inserirá a vinheta dramática que será apresentada logo a seguir, contendo um conceito ou conceitos do novo produto. Segue-se a

apresentação de uma seqüência de vinhetas introduzindo um enredo familiar que demonstra como o conceito do novo produto poderia ser usado.

Estes autores ressaltaram que os enredos familiares são derivados de seu trabalho com métodos etnográficos, que fundamentam o conceito do produto. A audiência participa, então, de algumas conversações estruturadas sobre o conceito introduzido, partindo da compreensão entremeada de implicações, questionamentos e expectativas quanto ao produto. Os projetistas do produto também participam das sessões, auxiliando, por um lado, o moderador na formulação das respostas aos questionamentos e, por outro, coletando idéias emitidas no decorrer das conversações.

Killingsworth *et al.* [Kill00] chamaram a atenção para as precauções a serem tomadas ao adotar as discussões em grupo, comentando que para que possam fornecer informações úteis é necessário considerar aportes válidos e efetivos. A seleção dos facilitadores (ou mediadores) e dos membros do grupo, bem como o planejamento adequado da duração e da pauta das sessões, constituem aspectos críticos para o sucesso da aplicação do método, devendo, pois, serem cuidadosamente ponderados.

Morgan [Morg93] e Krueger [Krue94] descreveram detalhadamente procedimentos metodológicos relativos às discussões em grupo, bem como aspectos relativos à obtenção de resultados satisfatórios, às implicações de adoção do método e vantagens e desvantagens de adotá-lo. Nielsen [Niel97b] discutiu aspectos relacionados ao uso e abuso ocorridos na prática das discussões em grupo. Dreachelin [Drea98] também oferece um suporte teórico com enfoque contextualizado para a condução de sessões de discussão em grupo, entremeando questões de cunho teórico inerentes ao método com implicações práticas de sua aplicação.

Assim como as entrevistas, as discussões em grupo podem ser úteis em processos avaliatórios formativos/evolutivos ou somativos/retrospectivos [McAt98]. Dentre os autores que relataram ensaios envolvendo esta técnica ou uma de suas variantes, podem ser mencionados McClelland e Brigham [McCl90], Yang [Yang90], McGraw [McGr94] e Sato [Sato99], Killingsworth *et al.* [Kill00].

3.2.1.9 Retorno de Opiniões do Usuário (*User Feedback*)/Ensaio de Usabilidade Remoto (*Remote Usability Testing*)

Estratégias de *retorno de opiniões do usuário*, também denominadas *ensaios de usabilidade remoto* (*remote usability testing*) por Hartson *et al.* [Hart96], requerem normalmente a conexão do sistema avaliado a uma rede de computadores e facilidades de correio eletrônico instaladas, podendo assumir diversas formas. Uma delas consiste em coletar reclamações e/ou elogios sobre um produto sob condições de teste através da disponibilização de endereços eletrônicos dedicados, "quadros de avisos" eletrônicos (*electronic bulletin boards*) ou "caixas de sugestões" *online* [Shne93]. Outras variantes incluem comandos especiais destinados à coleta automática de opiniões sobre as aplicações em teste ou a distribuição de versões beta do produto para que os usuários usem e dêem retorno de suas opiniões, sob a forma de relatórios *online*, verbalização por telefone ou relatos de caneta e papel, para a equipe de desenvolvimento daquele produto [Niel93b, Shne93, Popo95].

Consultoria online ou *por telefone* (*online* ou *telephone consultants*) é uma variante extremamente eficaz de prestação de assistência à distância, ao mesmo tempo que possibilita a coleta de informações sobre necessidades e problemas enfrentados pelo usuário, sugestões para otimização de facilidades e tendências potenciais de expansão de um produto [Shne93]. Diversos fabricantes disponibilizam um número para chamadas gratuitas a consultores e alguns sistemas em rede possibilitam a monitoração remota do terminal do usuário enquanto este mantém contato telefônico com um dos consultores.

Hiltz e Turoff [Hilt81] descreveram um sistema denominado *Electronic Information Exchange System - EIES* (*Sistema Eletrônico de Intercâmbio de Informações*), que permite aos usuários o envio de mensagens para uma caixa de correio eletrônico dedicada, denominada HELP, assim como o recebimento de respostas a tais mensagens, na maioria dos casos, em poucos minutos.

Caixa de sugestões online (*online suggestion box*) ou *relatórios de problemas* (*trouble reporting*) são variantes que envolvem facilidades de correio eletrônico, permitindo aos usuários o envio de mensagens via rede para fabricantes ou projetistas [Shne93]. Tais mensagens, sob a forma de sugestões e comentários construtivos ou de descrições de problemas enfrentados durante o uso de um produto particular, encorajam a participação do usuário no processo de aprimoramento do produto, ao mesmo tempo que fornecem informações de interesse à equipe de desenvolvimento. Shneiderman [Shne93] comenta que no *University of Maryland Computer Science Center* os usuários podem digitar o comando *GRIPE*⁴³ e receber a permissão de enviar comentários para o quadro de programação de sistemas.

Outra variante desta estratégia permite que uma comunidade usuária questione a aplicabilidade de um pacote de *software* a suas atividades cotidianas ou procure alguém com experiência de uso de tal sistema. Por não terem ninguém específico em mente, as facilidades de correio eletrônico não atendem a suas necessidades. Eis porque alguns projetistas de sistemas oferecem aos usuários um recurso que lhes permite enviar mensagens e convites "abertos", denominado *quadro eletrônico de avisos* (*electronic bulletin board*) [Shne93]. Muitos desses quadros de avisos contêm mensagens relativas a tópicos técnicos, e.g. linguagens de programação e problemas de *hardware*, enquanto outros são empregados como quadros de avisos convencionais.

Em suma, o *ensaio de usabilidade remoto* (*remote usability testing*) é um ensaio de usabilidade conduzido em situações nas quais os avaliadores e os usuários de teste se encontram separados espacial e temporalmente, impossibilitando a observação direta da interação usuário-computador pelos avaliadores ou a execução de tarefas de teste em um laboratório formal de usabilidade.

Hartson *et al.* [Hart96] revisaram várias formas de condução de ensaios remotos. Uma delas se aplica a situações em que avaliador e usuário de teste se encontram espacialmente separados, mas o monitoramento da interação usuário-computador pelo avaliador ocorre simultaneamente à execução das tarefas de teste pelo usuário. Neste

⁴³ Em inglês, reclamação, queixa, descontentamento.

caso, o avaliador monitora através de uma rede computacional o sistema no qual o usuário de teste executa as tarefas de teste. A comunicação entre ambos pode ocorrer via telefone ou recursos de videoconferência.

Há situações em que o avaliador, separado geograficamente do usuário de teste, não pode monitorar a execução das tarefas de teste à medida que estas ocorrem, o que exige outro modo de condução do ensaio remoto. Assim, o usuário de teste é guiado e monitorado por um aplicativo de *software* com o qual seu sistema é instrumentado, similantemente ao que ocorre em sessões de captura automática de dados, anteriormente discutidas.

A adoção deste método de avaliação da usabilidade exige, obviamente: (i) a conexão das máquinas do avaliador e do usuário de teste via rede de computadores, a partir da qual os resultados do ensaio são monitorados; (ii) o uso de aplicativos (e.g., *Look@me*) que possibilitem a observação, à distância, da tela do usuário de teste durante a realização do ensaio; (iii) recursos de videoconferência que permitam ao avaliador a comunicação com o usuário de teste (prestação de esclarecimentos, passagem de instruções, etc.) ou, no caso de se integrar a *verbalização de procedimentos (think aloud)* ao ensaio remoto, a comunicação do usuário de teste com o avaliador (descrição de ações durante a realização das tarefas de teste; e (iv) um aplicativo para a instrução do usuário de teste e a captura automática de dados (*automatic data logging*).

Em virtude da diversidade de partes dos sistemas terrestres multi-missões de recepção de dados e da variedade de grupos e projetos de pesquisa do *Jet Propulsion Laboratory* da *Caltech*, Elgin [Elgi95] ressaltou as dificuldades enfrentadas no desenvolvimento de um conjunto centralizado de facilidades de ajuda, bem como as limitações da adoção de muitos métodos de usabilidade a este contexto multivariado de subsistemas, configurações e usuários. Dada a variação da distribuição geográfica dos usuários dos produtos da *Caltech*, uma das alternativas viáveis apontadas pelo grupo de interesses específicos formado no CHI'95 para estudar estas questões foi o retorno de opiniões dos usuários via rede e a aplicação de ensaios de usabilidade remotos.

Pernice e Butler [Pern95] discutiram a aplicabilidade do *Lotus Notes*, uma plataforma cliente-servidor para desenvolvimento de aplicações de *groupware*, resultante da integração de um sistema de mensagens, um banco de dados de monitoramento de documentos e um ambiente de aplicação-desenvolvimento, com facilidades de: envio de fax para comunicação direta com usuários que não dispõem de acesso via *e-mail*, transmissão de vídeo em tempo real para terminais instalados em ambientes de trabalho de equipes de desenvolvimento de produtos a serem avaliados e recursos de videoconferência para a realização de ensaios de usabilidade remotos.

Um artigo da edição especial⁴⁴ de 01 de setembro de 1998 da revista *PC Magazine* [PCMO98] discutiu e comparou características de oito produtos empregados em ensaios de usabilidade remotos, a saber: *Compaq Carbon*, *Copy CoSession*, *Remote LapLink for Windows*, *NetOp for Windows*, *pcAnywhere32*, *RapidRemote*, *ReachOut Enterprise* e *Remotely Possible*. Esta revisão é complementada pela revisão informal realizada por Preston [Pres99], que ressaltou que o termo *ensaio remoto* em toda a literatura por ela revisada tende a ser aplicado ao tipo de

⁴⁴ *Remote-Control Software (Software para Controle Remoto)*.

ensaio que envolve o monitoramento em tempo real das atividades de teste realizadas em um sistema computacional remoto.

Através de praticamente todos os produtos revisados os participantes formulam questões "ao vivo" (via conexão telefônica, Internet ou sessões de diálogo digitado), embora Preston [Pres99] mencione a existência de alguns produtos (e.g, *ErgoLight Lab Tester* [EL00]) que possibilitam a condução de ensaios remotos sem necessidade de monitoramento em tempo real por parte do avaliador e acrescenta que muitos dos produtos possuem facilidades de registro e/ou *playback*, que atuam como uma extensão natural do monitoramento remoto. Outras denominações encontradas por Preston para produtos desta natureza são *software para controle remoto*, *ferramenta de telecomunicação (telecommuting tool)*, *ferramenta para monitoramento de sistemas (system administration tool)* e *software para videodiálogo (video chat software)*.

Mayhew [Mayh99] descreve como formas alternativas de condução da avaliação de modelos conceituais iterativos, quatro das técnicas discutidas por Hartson *et al.* [Hart96] – *avaliação por controle remoto (remote-control evaluation)*, *videoconferência (video conferencing)*, *avaliação remota instrumentada (instrumented remote evaluation)* e *avaliação remota semi-instrumentada (semi-instrumented remote evaluation)*.

Nielsen [Niel00] comenta que os testes através da Web podem ser realizados em nível internacional com uma grande economia de custos de viagem. Como os usuários têm acesso a quase todo o mundo via Web, estes podem acessar o *site* a ser testado. O autor sugere, assim como o fizeram Hartson *et al.* [Hart96], a comunicação telefônica convencional ou o uso de facilidades de vídeoconferência, embora externamente sua preferência por diálogos telefônicos, argumentando que atividades de vídeoconferência exigem recursos extras de *hardware* e *software*, além de dividirem a atenção do avaliador entre o recrutamento do usuário e os ajustes do sistema. No entanto, Nielsen [Niel00] chama a atenção para a questão das barreiras idiomáticas quando se opta por contatos telefônicos convencionais.

Uma variante do *ensaio remoto* apontada por Nielsen [Niel00] é o *ensaio auto-administrado (self-administered test)*, onde as instruções dos testes são dadas por telefone ao usuário que, posteriormente, relata os resultados ao avaliador via correio convencional, *e-mail* ou formulários *online*.

Hill e Terveen [Hill97] desenvolveram *PHOAKS*, um sistema de reconhecimento automático de URLs recomendadas em mensagens da Usenet e a atualização contínua de um *site* de Web no qual os dados coletados remotamente são sintetizados. As páginas automaticamente geradas são visualizadas como "esboços inacabados" que os usuários auxiliam a refinar.

Segundo Castillo *et al.* [Cast97a, Cast97b, Cast98] e Hartson e Castillo [Hart98], os tipos usuais de testes alfa e beta não se qualificam como processos avaliatórios formativos, por não produzirem dados quer detalhados durante a observação do uso, quer estreitamente relacionados com o desempenho de tarefas específicas. Além disso, argumentaram que a identificação de incidentes críticos é a única e a mais importante fonte de dados desta natureza. Partindo desta premissa, desenvolveram e validaram um método econômico de avaliação, fundamentado no auto-relato de incidentes críticos por usuários reais, a partir da execução de tarefas reais em seus

ambientes de trabalho.

Conforme os resultados obtidos, Hartson e Castillo [Hart98] afirmam que com apenas um breve treinamento, os usuários são capazes de identificar, relatar e ponderar o grau de severidade de seus próprios incidentes críticos. Participaram voluntariamente do estudo 24 estudantes (6 do sexo feminino e 18 do sexo masculino, 22 estudantes de graduação de Ciência da Computação e 2 graduados), selecionados a partir de um questionário de sondagem do perfil dos usuários de teste segundo o critério do menor conhecimento de busca e aquisição de informações disponíveis na Web. O aplicativo avaliado no estudo foi a *Internet Movie Database (IMDb)*, disponível em www.imdb.com. Este é um serviço que oferece acesso livre e extensivo a informações sobre filmes, enquanto uma ferramenta de relato de incidentes críticos permite aos usuários o envio de relatórios estruturados sobre os incidentes críticos identificados ao longo das sessões experimentais.

Zhang *et al.* [Zhan98a] descreveram um estudo comparativo da usabilidade de três alternativas para a comunicação de voz em conjunção com serviços da Web para atendimento ao consumidor. Duas das tecnologias usaram uma única linha telefônica para a transmissão de voz e dados – os serviços telefônicos para a Internet e a transmissão simultânea de voz e dados (*Simultaneous Voice and Data - SVD*), um protocolo destinado ao envio de sinais de voz através de uma rede telefônica pública, ao invés do envio através da Internet. A terceira alternativa empregou uma segunda linha telefônica para o contato com o consumidor. Neste experimento, os autores relataram o uso de facilidades de avaliação remota da usabilidade.

Scholtz [Scho98] conduziu um ensaio remoto baseado em quiosque⁴⁵ visando atingir simultaneamente duas metas-chaves, uma da equipe de desenvolvimento e outra da equipe de pesquisa da usabilidade. A primeira meta era relativa à avaliação da usabilidade do projeto de um manual técnico *online* (disponível na Web), enquanto a segunda dizia respeito à determinação da factibilidade do uso de quiosques de usabilidade para a coleta de dados referentes a questões de projeto de livros *online*, bem como a identificação das vantagens e desvantagens associadas à estratégia planejada. Os testes foram conduzidos duas vezes, a primeira monitorada pela equipe de avaliação, a última pela equipe de projeto, com o propósito de confrontação de resultados.

Nodder *et al.* [Nodd99] reportaram iniciativas de avaliação da usabilidade do *NetMeeting* da *Microsoft* em suas diferentes versões, destacando aspectos relativos às mudanças ocorridas nas categorias de usuários, na natureza das tarefas e nos métodos de avaliação empregados ao longo da evolução das versões do produto. Os autores mencionaram que as descobertas feitas durante as avaliações do *NetMeeting* envolvendo atividades de vídeoconferência possibilitaram encorajar os demais membros da equipe de avaliação da *Microsoft* a usarem o produto em ensaios de usabilidade remotos.

⁴⁵ *Ensaio baseado em quiosques (Kiosk-based testing)* são ensaios planejados para serem realizados em quiosques disponíveis a usuários voluntários em locais públicos (e.g., bancos 24 horas), sem nenhum auxílio humano direto. Estes ensaios caracterizam-se por tarefas completamente auto-explicativas de curta duração, mas que possibilitem a coleta de dados significativos ao contexto.

3.2.2 Inspeções de Usabilidade (*Usability Inspections*)

Em geral, *inspeções de usabilidade* são estratégias avaliatórias fundamentadas na análise e julgamento de projetos por avaliadores (ergonomistas, engenheiros de *software*, engenheiros de usabilidade), que investigam aspectos relativos à usabilidade segundo um conjunto de critérios, recomendações, normas ou heurísticas. Nielsen e Mack [Niel94a] também inserem neste contexto usuários finais com conhecimento do contexto ou da tarefa avaliada, assim como outros profissionais com habilidades de avaliadores, embora Constantine [Cons94] trate as estratégias de inspeção envolvendo usuários como *inspeções colaborativas de usabilidade (collaborative usability inspections)*.

Os diferentes métodos de inspeção da usabilidade visam a avaliação de projetos de interfaces usuário-computador, variando conforme as metas almejadas, a forma a partir da qual o julgamento é emitido e os critérios nos quais os avaliadores fundamentam seus pareceres [Niel94a]. Segundo Constantine e Lockwood [Cons99], pode-se inspecionar quase tudo, desde documentos de especificação a protótipos de telas, de modelos de projetos a protótipos em papel, de sistemas em funcionamento a protótipos funcionais. Dix *et al.* [Dix98] são mais ortodoxos neste aspecto, associando os métodos de inspeção a processos avaliatórios de projetos e sendo explícitos ao acrescentarem “antes da implementação”.

Diversos autores descreveram métodos de inspeção, quer como procedimentos emergentes ou aplicados especificamente (e.g., Tetzlaff [Tetz91], Lewis e Polson [Lew91b], Dzida [Dzid96]), quer inseridos no contexto mais abrangente dos processos interativos usuário-computador ou com um caráter acadêmico (e.g., Nielsen e Mack [Niel94a], Dix *et al.* [Dix98], Constantine e Lockwood [Cons99]). Alguns, e.g., Nielsen e Mack [Niel94a] e Dix *et al.* [Dix98], incluem estratégias de avaliação formais na coletânea de métodos de inspeção descritos, as quais envolvem aspectos de usabilidade relativos ao código ou aos requisitos de projetos. Este documento, no entanto, contém apenas métodos de inspeção passíveis de aplicação na avaliação da implementação, tendo em vista que a pesquisa aqui documentada é centrada na avaliação de produtos finais.

Assim, as estratégias de inspeção da usabilidade discutidas neste capítulo serão quatro, a saber:

- ⊙ Revisões Sistemáticas (*Walkthroughs*);
- ⊙ Inspeção Baseada em Diretrizes de Projeto, Guias de Estilo e Padrões (*Design Guidelines, Styleguides and Standards Inspection*);
- ⊙ Avaliação Heurística (*Heuristic Evaluation*);
- ⊙ Inspeção Fundamentada na Perspectiva (*Perspective-Based Inspection*); e

3.2.2.1 Revisões Sistemáticas (*Walkthroughs*)

O conceito básico de *revisão sistemática (walkthrough)* não é recente, nem tampouco específico da área de engenharia de *software*. Trata-se simplesmente da análise estruturada de qualquer produto por uma equipe de especialistas [Your79, Riem91]. No contexto da engenharia de *software* o produto analisado poderá ser um programa ou um sistema computacional. No domínio dos

processos interativos homem-sistema computacional, o enfoque dado ao produto avaliado recai freqüentemente em sua interface usuário-computador, tendo em vista que a comunicação do usuário com o produto se dará através dessa interface.

Como quaisquer outros procedimentos de análise, as *revisões sistemáticas* podem ocorrer ao longo das diferentes etapas de desenvolvimento de um sistema, assumindo conotações e formas variadas e envolvendo competências as mais diversas, embora seus fundamentos se mantenham inalterados: um grupo de especialistas - indivíduos desempenhando funções de *status* equiparados numa organização ou em organizações afins - reunidos para revisar e discutir um produto específico.

Segundo Yourdon [Your79], para a maioria das aplicações disponíveis no mercado nos dias atuais, o indicador de *qualidade* mais importante é a *correção* (*correctness*), sendo embaraçoso, caro e, às vezes, desastroso, lançar no mercado um produto que ainda contenha falhas. Além do mais, estratégias clássicas de "testes exaustivos" falham quando se trata de descobrir falhas em sistemas computacionais complexos como os sistemas atuais. Por outro lado, as *revisões sistemáticas* têm se mostrado altamente satisfatórias em atividades de produção de sistemas confiáveis e livres de falhas. Grupos de desenvolvimento trabalhando com *revisões sistemáticas* têm relatado uma redução de até dez vezes no número de erros na produção de aplicativos, quando comparados com procedimentos convencionais.

Yourdon distingue quatro categorias de *revisões sistemáticas*: (i) *revisões sistemáticas de especificação*, que consistem da análises de requisitos funcionais ou especificações de um sistema computacional, envolvendo normalmente um analista de sistemas, um usuário representativo e um ou mais projetistas, com o propósito de focalizar problemas, imprecisões, ambigüidades e omissões cometidos no processo de especificação de uma aplicação computacional; (ii) *revisões sistemáticas de projeto*, que enfatizam soluções de problemas de natureza lógica, física ou procedural, considerando que os requisitos funcionais do sistema foram corretamente especificados; (iii) *revisões sistemáticas de código*, cujo alvo é o código de uma dada aplicação embora, às vezes, infelizmente não verifiquem problemas de análise ou de projeto, para a consternação do programador, que dispendeu dias escrevendo um código elegante para a aplicação; e (iv) *revisões sistemáticas de teste*, conduzidas para garantir mais a adequação dos dados de teste à aplicação do que para examinar a saída do teste propriamente dita, e realizadas tipicamente por uma equipe composta pelo indivíduo que produziu os dados de teste, programadores, um analista de sistemas e um usuário representativo.

Uma revisão sistemática bem estruturada envolve uma equipe de indivíduos, cada um dos quais desempenha um papel definido, embora não necessariamente definitivo. Um indivíduo pode exercer a função de *coordenador* em uma revisão sistemática e de *secretário* em outra posterior ou até mesmo acumular funções em um único processo.

No âmbito dos processos interativos, a estratégia foi combinada com um modelo cognitivo de aprendizado por exploração, resultando na *revisão sistemática cognitiva* (*cognitive walkthrough*), onde um grupo de engenheiros avalia discrepâncias entre as metas e expectativas dos usuários e os procedimentos requeridos pela aplicação [Lewi90, Lewi91b, Riem91, Mill92, Pols92].

Através desta metodologia, projetistas de interfaces "passeiam" através de um projeto de prancheta no contexto de tarefas-chaves que um usuário típico deverá executar, comparando as ações da interface com as metas e o conhecimento do usuário, observando as discrepâncias entre as expectativas do usuário e o comportamento real da interface, repassando e questionando os mais diferentes aspectos do projeto, tais como seqüenciamento lógico de telas e da informação apresentada pela interface, a usabilidade de cada tela e a consistência do projeto em diferentes níveis [Jeff91, Bias91].

Deste modo, torna-se possível a avaliação da usabilidade de um sistema nos estágios iniciais do processo de projeto [Riem91, Niel93c]. Embora não seja infalível nem permita antecipar soluções para determinados problemas, trata-se de uma estratégia valiosa por permitir (i) prever satisfatoriamente grande parte do conhecimento requerido para solucionar problemas [Bell91]; (ii) auxiliar a descobrir pontos carentes de otimização em interfaces existentes e (iii) criar novas interfaces [Habe91].

Rieman *et al.* [Riem91] conceberam a *revisão sistemática cognitiva automatizada (automated cognitive walkthrough)*, uma variante na qual o sistema substitui os formulários impressos que constituem o maior inconveniente do método. Muitas das respostas a questões relativas a cada uma das ações⁴⁶ do usuário em tarefas específicas, que são tipicamente registradas em formulários por um analista, podem ser coletadas a partir de um simples acionamento do *mouse* ou da entrada de um texto curto pelo teclado, dispensando a tediosa tarefa de preenchimento de páginas e páginas de perguntas.

Vale a pena ressaltar que a metodologia concebida por Lewis *et al.* [Lewi90] destinava-se originalmente à avaliação de interfaces simples do tipo *walk and use*⁴⁷. Este fato é reforçado na documentação de aplicações do método, e.g. Lewis e Polson [Lewi91b], Rieman *et al.* [Riem91], Polson *et al.* [Pols92], Rieman *et al.* [Riem95].

Franzke [Fran91], Jeffries *et al.* [Jeff91] e Lewis e Polson [Lewi91b], estimulados pelas restrições de aplicação da proposição original de Lewis *et al.* [Lewi90] apenas a interfaces do tipo *walk and use*, iniciaram investigações da extensão do método a contextos de avaliação de interfaces de ambientes mais complexos, mas ainda de suporte a usuários ocasionais ou principiantes. Durante o desenvolvimento destas iniciativas, Franzke [Fran91] e Jeffries *et al.* [Jeff91] definiram como uma das metas da investigação, a condução das avaliações em contextos mais realistas, envolvendo ambientes industriais e grupos de projetistas de *software*, ao invés de especialistas de interação usuário-computador.

No estudo comparativo envolvendo quatro métodos de avaliação da usabilidade, Jeffries *et al.* [Jeff91] executaram algumas alterações na *revisão sistemática cognitiva* proposta por Lewis *et al.* [Lewi90], visando uma adaptação ao contexto do experimento, a saber: (i) participação de uma

⁴⁶ Os autores empregam o termo *ações* no sentido de atitudes em nível do acionamento de teclas (*keystroke*) ou da seleção de itens em menus.

⁴⁷ *Passeie e use*, traduzindo literalmente para português. Trata-se de interfaces dotadas de mecanismos de interação usualmente fundamentados em diálogos via menus ou manipulação direta, que apresentam listas de itens textuais ou barras de ferramentas contendo itens gráficos, sobre os quais o usuário "passeia", "usando" os itens de seu interesse, a partir de seleções via *mouse* ou dispositivo equivalente.

equipe de três engenheiros de *software* dos *HP Laboratories*, a fim de tornar a estratégia mais consistente com os procedimentos usuais adotados em revisões sistemáticas de *software*; (ii) adoção de um procedimento adaptativo de captura das considerações dos diversos tipos de usuários, especificado por um dos investigadores do trabalho original de revisão sistemática cognitiva, com o propósito de atender aos requisitos da interface de teste, do tipo "localiza e usa" (*walk up and use*); (iii) fundamentação em tarefas, embora sem indicação de *como* acessá-las; e (iv) refinamento do procedimento e das tarefas a partir de um experimento-piloto prévio envolvendo revisão sistemática cognitiva.

Bias [Bias91] denominou *revisão sistemática de usabilidade* (*usability walkthrough*) o procedimento de revisão sistemática adaptado para a avaliação da usabilidade. Segundo Bias, trata-se de uma revisão sistemática de um projeto ainda no papel que responde a inúmeras questões relativas ao encaminhamento lógico e à consistência de um determinado projeto. O autor descreveu sua experiência com testes de usabilidade na *IBM Corp.* e a variante produzida pela equipe de desenvolvimento desta organização, denominada *revisão sistemática pluralista de usabilidade* (*pluralistic usability walkthrough*). A estratégia concebida pela equipe da *IBM-Austin* apresenta as seguintes características: (i) participação de três categorias de indivíduos - representantes da população usuária, projetistas do produto e profissionais de fatores humanos; (ii) apresentação dos painéis da interface de usuário na mesma seqüência em que são confrontados online; (iii) antecipação das respostas dos participantes (sobre questões relativas a cada um dos painéis) a qualquer discussão suscitada pelo tópico em questão; e (iv) antecipação da verbalização dos usuários representativos, estimulada até a exaustão dos comentários de cada usuário, à verbalização de questões ou dúvidas remanescentes da equipe de especialistas.

Bell *et al.* [Bell91] adaptaram o procedimento de revisão sistemática visando a avaliação da facilidade de projeto de uma linguagem gráfica de programação antes de sua implementação. O procedimento adaptado foi denominado *revisão sistemática de programação* (*programming walkthrough*) e consistiu de uma análise, em nível do conhecimento, da linguagem *ChemTrains*. Destinada à criação de simulações animadas por usuários sem fundamentação de programação, a estratégia visou a identificação de aspectos específicos necessários ao bom desempenho de uma ou mais tarefas pré-definidas, fundamentadas na linguagem gráfica de programação *ChemTrains*.

A partir dos estágios mentais que um programador deveria adotar para solucionar problemas específicos, o método de Bell e colegas permite a avaliação de dois aspectos importantes de um ambiente de programação - *expressividade* (*expressiveness*), capacidade de encontrar, de modo simples, soluções para problemas; e *facilidade* (*facility*), capacidade de solucionar problemas facilmente. Conseqüente de procedimentos de revisão sistemática cognitiva, a *revisão sistemática de programação* requer um conjunto representativo de tarefas ou problemas e um documento descrevendo as necessidades de um usuário principiante no tocante ao conhecimento do sistema⁴⁸.

Wharton *et al.* [Whar92] discutiram as questões abordadas e as recomendações propostas por Franzke [Fran91] e Jeffries *et al.* [Jeff91] para a aplicação de revisões *sistemáticas cognitivas*

⁴⁸ Tal documento é comumente denominado *dogma* (*doctrine*) e inclui conceitos gerais do sistema e seu uso, além de recomendações sobre como confrontar e solucionar problemas.

em contextos realísticos, descrevendo (i) ensaios não controlados com três interfaces com o usuário complexas, (ii) problemas recorrentes observados e (iii) respectivas repercussões sobre o uso do método em outros contextos avaliatórios e sua condução por outros avaliadores.

Rowley e Rhoades [Rowl92], por outro lado, focalizaram o dispêndio de tempo na condução de revisões sistemáticas, assim como a impopularidade do método junto a avaliadores investigando problemas de usabilidade associados a tarefas com grau de complexidade elevado. Visando a maximização das informações úteis obtidas a partir de revisões sistemáticas, minimizando simultaneamente os esforços associados ao procedimento, Rowley e Rhoades propuseram uma variante para a revisão sistemática, a ser conduzida a "um ritmo mais acelerado"⁴⁹, o *jogthrough*. Incorporando registro em vídeo a uma sessão de revisão sistemática, o método foi validado em meio às restrições impostas por um ambiente real de desenvolvimento de produtos, tendo revelado problemas significativos na interface avaliada, cuja análise poderia, então, ser refinada a partir de outros métodos avaliatórios.

Com o propósito de compensar os problemas associados ao tempo de condução da revisão sistemática, Rowley e Rhoades [Rowl92] alteraram o procedimento tradicional registrando a sessão de avaliação em vídeo ao invés de transcreverem os comentários manualmente. Empregando um aplicativo anteriormente desenvolvido para a captura de dados relativos a sessões de avaliação da usabilidade com registro de vídeo, os autores foram capazes de capturar eventos significativos em tempo real na sessão de revisão sistemática adaptada. Neste contexto, os papéis dos participantes de revisões típicas são alterados, o membro responsável pelas anotações torna-se um operador da câmara de vídeo e do aplicativo de captura de dados. O aplicativo de captura de dados permitiu o registro de observações referentes aos eventos da avaliação, cada um dos quais foi rotulado com informações temporais, a partir dos recursos da câmara de vídeo.

Durante a sessão de *jogthrough*, surgiram diversas sugestões relevantes de projeto, que teriam sido suprimidas pelo moderador, caso se tratasse de uma revisão sistemática típica. No entanto, o procedimento modificado permitiu a formulação de sugestões apropriadas, a fim de explorar caminhos alternativos para o usuário em função de suas escolhas ao longo do ensaio. Para tanto, foi necessária uma alteração das regras básicas da revisão típica, com o objetivo de permitir uma discussão mais livre, menos restrita aos questionamentos contidos no plano de avaliação, à medida que foram percebidas no diálogo tendências a sugestões relevantes ao contexto analisado.

Obviamente, tal grau de liberdade no critério, segundo o qual o moderador julga a propriedade das discussões para o contexto da avaliação, torna o procedimento mais subjetivo. Portanto, a seleção de tarefas torna-se uma questão de grande importância no contexto do *jogthrough*. Na opinião de Rowley e Rhoades [Rowl92], em algumas circunstâncias, é necessário negligenciar aspectos que poderão conduzir à descobertas de problemas de usabilidade apresentados pelo produto avaliado. Apesar de serem decisões nem sempre fáceis de serem tomadas, o procedimento de *jogthrough* fornece um

⁴⁹ Em português, não se consegue fazer o trocadilho feito em inglês entre *to walk* (caminhar) e *to jog* (sacudir, trotar, acelerar o passo), dando a idéia de que o *jogthrough* consiste em uma "caminhada em ritmo acelerado", i.e de que o *jogthrough* seria uma versão abreviado do *walkthrough*.

mecanismo de identificação de aspectos da interface que requerem uma atenção extra, permitindo aos avaliadores a identificação desses aspectos de modo mais rápido do que ocorre no caso de se adotar uma revisão sistemática típica.

Após quatro anos de amadurecimento da metodologia original⁵⁰, Wharton *et al.* [Whar94] apresentaram uma segunda revisão das metas originais propostas para a *revisão sistemática cognitiva*, transferindo a ênfase da consideração explícita da estrutura de metas do usuário para a motivação da escolha de ações corretas pelo usuário a partir de um embasamento em teoria cognitiva. Segundo os autores, o procedimento original foi concebido como um simples processo de pergunta e resposta, que avaliadores (projetistas) com pouca fundamentação em teoria cognitiva poderiam utilizar, enquanto esta terceira versão incorporava a constatação da necessidade de uma compreensão básica da teoria cognitiva e, a partir da qual seria estruturado o processo de revisão para melhor se adequar às necessidades impostas por situações particulares.

McGraw [McGr94] descreveu a combinação de discussões em grupo (*focus groups*) com revisões sistemáticas, na etapa de *criação e avaliação de quadros de eventos* (*storyboards*) do processo de desenvolvimento de aplicações, centrado no usuário, que a *Cognitive Technologies* adota. A revisão sistemática é realizada em nível de quadro de eventos, num contexto mais amplo em que um conjunto de quadros de eventos são distribuídos sobre expositores ao longo de uma sala de reuniões. Cada quadro de eventos é projetado em grandes dimensões numa tela de projeção, enquanto a equipe de revisão sistemática dá aos participantes uma visão analítica pluralista da situação. Após as atividades de revisão terem sido concluídas e os usuários de teste terem passeado diante da seqüência de quadros de eventos sob avaliação, fazendo anotações/comentários e respondendo a questões relativas a cada contexto analisado, segue-se um procedimento de discussão em grupo e o fechamento do processo avaliatório daquela etapa.

Ereback e Höök [Ereb94] conduziram uma revisão sistemática cognitiva na avaliação da interface de um sistema de reserva de inscrições, retomando à linha de investigação de Franzke [Fran91], Jeffries *et al.* [Jeff91] e Lewis e Polson [Lewi91b] sobre a aplicabilidade do método a procedimentos avaliatórios de aplicações mais complexas, neste caso envolvendo a colaboração suportada por computadores (CSCW). Os autores revisaram a metodologia proposta por Polson *et al.* [Pols92], assim como os estudos comparativos desenvolvidos por Jeffries *et al.* [Jeff91] e Karat *et al.* [Kara92] e o estudo de aplicação do método a interfaces mais complexas, conduzido por Wharton *et al.* [Whar92], visando estender as investigações anteriores às aplicações CSCW. Os autores concluíram que embora útil ao contexto, o método carecia de alterações, dentre as quais foram mencionadas (i) a expansão da teoria subjacente ao método, visto que as metas estabelecidas para aplicações mono-usuário sofrem freqüentemente a influência de outros usuários, típica de ambientes multi-usuários, e.g. CSCW; e (ii) a necessidade de enriquecimento das descrições dos usuários, visto que o envolvimento de vários usuários torna a condução da revisão sistemática mais difícil do que quando é conduzida com um único usuário.

⁵⁰ Descrita por Wharton *et al.* [Whar92].

John e Packer [John95a] descreveram um estudo de caso do aprendizado e condução de *revisões sistemáticas cognitivas (cognitive walkthroughs)* por projetistas, no processo avaliatório de uma ferramenta multimídia de aprendizado, apresentando recomendações úteis para projetistas de sistemas e pesquisadores de métodos de avaliação. Conforme os autores, embora o método seja usável e de aprendizado relativamente fácil, há vários aspectos a serem investigados a fim de otimizar o uso do método por projetistas de *software*. Além disso, a investigação de John e Packer delinea um cenário realístico que pode servir de respaldo para análises ulteriores envolvendo o aprendizado e uso do método por avaliadores principiantes.

Nielsen [Niel95], ao descrever o plano de usabilidade adotado pela equipe de desenvolvimento da interface de usuário da *SunWeb*, o sistema de informações WWW interno da *Sun Microsystems*, mencionou *revisões sistemáticas* dentre as diferentes estratégias empregadas no processo avaliatório. Segundo o autor, *revisões sistemáticas* foram incluídas como ensaios finais de usabilidade. A fim de evitar o problema de ativação de teclas ainda não associadas a algumas das ações do sistema, preferiu-se uma cópia colorida ampliada da tela do projeto da *home page* da *SunWeb* ao invés da própria tela. A revisão sistemática envolveu a solicitação aos usuários de teste da associação de cada tecla à informação que eles imaginavam poder acessar através de seu acionamento, bem como uma descrição de cada ação praticada. Ao final da revisão, solicitou-se de cada usuário de teste comentários sobre a estética dos ícones, uma listagem dos ícones, e listagens daqueles mais e menos apreciados no contexto da interface.

Rizzo *et al.* [Rizz97] apresentaram, no âmbito do projeto *AVANTI*⁵¹, uma variante da *revisão sistemática cognitiva* proposta por Lewis *et al.* [Lewi90]. De acordo com os autores, a diferença entre o método que conceberam e o método de Lewis *et al.* [Lewi90] reside, em essência, na fundamentação de ambos. Enquanto o método de Rizzo *et al.* [Rizz97] se fundamenta no modelo de ação humana desenvolvido por Norman [in Hutc85], o método de Lewis *et al.* [Lewi90] se baseia na teoria de aprendizagem exploratória de Polson e Lewis [Pols90] e no modelo GOMS de Kieras e Polson [Kier85, Kier88].

Em linhas gerais, a *revisão sistemática cognitiva* proposta por Rizzo *et al.* [Rizz97] investiga as atividades do usuário relacionadas com uma ou mais tarefas contextualizadas. O avaliador investiga o sistema selecionando ações que possam contribuir para a execução da tarefa, i.e., ações cuja descrição ou aparência seja compatível com o que se intenta realizar. Após a seleção das ações pertinentes ao contexto da tarefa, segue-se a interpretação da resposta do sistema, com o fim de avaliar se o progresso deu-se no sentido de conclusão da tarefa ou de reconsideração da meta. Deste modo, torna-se possível identificar se o significado e o estilo da interface foram bem interpretados pelo usuário, assim como se este é capaz de atingir metas factíveis e executar a ação correta sobre o objeto correto. De acordo com os roteiros definidos junto aos usuários, cada membro dos grupos de projeto e avaliação executa uma tarefa específica, formulando, a

⁵¹ *AVANTI (AdaptiVe and Adaptable iNteractions for multimedia Telecommunication applications)* é um projeto do ACTS (*Advanced Communication Technology & Services*), um programa europeu voltado para os problemas relativos (i) à aplicação das telecomunicações na difusão de informações multimídia para diferentes categorias de usuários, em especial deficientes físicos e idosos; (ii) à implantação de aplicações experimentais; e (iii) à experimentação de resultados obtidos em trabalhos de campo.

cada passo da interação, as questões da lista apresentada no Quadro 12.

Quadro 12 –Lista de questões empregada na *revisão sistemática cognitiva* proposta por Rizzi *et al.* [Rizz97]

LISTA DE QUESTÕES	
Q1:	A ação (ou o conjunto de ações) possível e correta será suficientemente evidente para o usuário e, se o for, é compatível com a sua intenção? (Intenção-Ação)
Q2:	O usuário consegue estabelecer a conexão entre a descrição da ação correta e o que ele intenta fazer? (Ação- Forma)
Q3:	O usuário receberá um retorno no mesmo local e com o mesmo estilo nos quais ele executou a ação? (Entrada: Ação- Saída: Retorno de Informação)
Q4:	O usuário interpretará corretamente a resposta do sistema para a ação executada, i.e., o usuário saberá se fez a escolha correta a partir da resposta do sistema? (Resultado- Forma)
Q5:	O usuário analisará corretamente os resultados, i.e., o usuário conseguirá verificar se está mais próximo de sua meta? (Forma-Análise)
Q6:	Se a meta tentada não for exequível (ou puder ser reformulada/otimizada), o usuário conseguirá entender que o que ele está tentando realizar não pode ser concluído nas condições atuais (ou o usuário encontrará outras metas? (Ação Futura-Foco)

Conforme Rizzo *et al.* [Rizz97], para qualquer resposta não inteiramente afirmativa, a questão é comunicada aos demais participantes, juntamente com a especificação das soluções alternativas que serão *a posteriori* implementadas e repetidamente testadas até que se obtenha uma resposta completamente afirmativa.

Sears e Hess [Sear98] retomaram a linha de pesquisa de Wharton *et al.* [Whar94] no sentido de incrementar o estudo revisivo do procedimento de *revisão sistemática cognitiva* com o propósito de torná-lo mais acessível aos avaliadores, além de otimizá-lo no tocante ao tempo dispendido na aplicação do método. Sears e Hess se concentraram mais especificamente na redução do tempo de treinamento formal em psicologia cognitiva, focalizando para tanto a concepção de soluções passo-a-passo detalhadas para as tarefas investigadas.

Tendo constatado que tais soluções detalhadas eram dependentes de diversos fatores, dentre os quais a experiência dos avaliadores com a condução do procedimento, os autores exploraram o impacto do detalhamento da descrição das tarefas no desempenho de avaliadores recém-treinados, tendo concluído que pequenas alterações podiam ter efeitos significativos sobre indivíduos ainda em fase de aprendizado da aplicação de novos métodos de avaliação. Outra consequência importante da pesquisa de Sears e Hess foi a evidência de que o detalhamento nas descrições das tarefas podia alterar o modo como os avaliadores aplicavam a *revisão sistemática cognitiva*.

Diversos autores relataram ensaios envolvendo este método ou uma de suas variantes, e.g., Cuomo e Bowen [Cuom92], Franzke [Fran95], Frederickson-Mele e Conrad [Fred99] e Novick [Novi99].

3.2.2.2 Inspeção Baseada em Diretrizes de Projeto, Guias de Estilo e Padrões (*Design Guidelines, Styleguides and Standards Inspection*)

Um meio popular para a difusão de conhecimento sobre fatores humanos tem sido a compilação de documentos contendo regras gerais ou específicas de projeto de interfaces, sendo as formas mais comuns de apresentação as *diretrizes* e os *guias de estilo* (*styleguides*) [Sond82, Löwg92]. Via de regra, as *diretrizes de projeto* se apresentam sob a forma de sugestões e recomendações técnicas que sumarizam o conhecimento e a opinião metodológica corrente [Krue83], traduzindo princípios bem conhecidos aplicáveis a projetos de interfaces de usuário em desenvolvimento [Niel93b]. Por outro lado, os *guias de estilo*, documentam padrões industriais e contêm informações prescritivas, ao invés de sugestivas, resumindo normas relacionadas com o estado da arte dos dispositivos interativos disponíveis no contexto considerado.

Além do mais, tais compêndios de regras, recomendações e diretrizes, embora constituam uma das ferramentas mais populares para a definição de um *estilo* (*look and feel*) de projeto, "atrelam" os esforços dos projetistas a padrões industriais e organizacionais e direcionam o conceito de *consistência* para plataformas ou linhas de produtos específicas [Root93]. No contexto da avaliação de interfaces, o uso de documentos dessa natureza (e.g., [Gali85, Shne87, Mayh92, Quei94, Micr95, Appl97]) respaldam os avaliadores com recomendações das mais diversas naturezas sobre o projeto de uma interface sob condições de teste, e.g., como as diferentes componentes do diálogo devem ser organizadas ou como os itens de menus deveriam ser dispostos [Jeff91, Treu94].

Entretanto, Constantine e Lockwood [Cons99] alertaram para o fato de tanto as recomendações dos *guias de estilo* quanto os *padrões de interfaces* serem instrumentos auxiliares no projeto (e avaliação) de interfaces que carecem de revisão contínua, a ser determinada pelos resultados de atividades de modelagem e projeto. Tal esforço assegurará a validade, a relevância e a atualização do conteúdo de compilações desta natureza, resultando em um processo contínuo e adaptativo da documentação ao invés de prescrições rígidas. É conveniente observar que tal advertência também se aplica às *diretrizes de projeto*, pois mesmo contendo informações sugestivas, tais informações são dependentes de avanços tecnológicos (e.g., novos dispositivos de entrada e/ou de saída), carecendo ser revisadas e atualizadas continuamente.

É conveniente salientar que tanto as coletâneas de *diretrizes de projeto* quanto os *guias de estilo* são formas de apresentação de conhecimentos e experiências destinadas à vasta gama de indivíduos atuantes no domínio da interação homem-máquina, desde estudantes e novos projetistas até especialistas e avaliadores de algoritmos. todos podem se beneficiar do uso de diretrizes de projeto e guias de estilo tanto como ponto de partida para a implementação de novas aplicações baseadas no grau de satisfação dos usuários [Bail83], quanto como fio condutor para a concepção de produtos usáveis e consistentes, de conformidade com convenções consagradas [Tetz91].

As *diretrizes de projeto* variam em conteúdo, podendo ser categorizadas como *genéricas*, quando enumeram princípios gerais e se aplicam a qualquer contexto de projeto;

ou, por outro lado, *específicas*, quando explicitam detalhes particulares a contextos [Tetz91, Niel93b]. As *diretrizes específicas* ainda podem ser subdivididas em *específicas a categorias*, quando enumeram detalhes sobre determinadas classes de produtos (e.g., sistemas baseados em janelas ou sistemas com interfaces de reconhecimento de voz) ou *específicas a produtos*, quando aplicáveis a uma linha de produtos particular (e.g., aplicações específicas ao Windows) [Niel93b]. Algumas coletâneas focalizam a apresentação e o comportamento da interface; outras dizem respeito a características de desempenho da interface em uso [Tetz91] ou encerram recomendações gerais de usabilidade [Shne87, Shne93].

Diferentes grupos de desenvolvimento de ferramentas de *software*, quer de caráter aplicativo específico ou de caráter criativo, têm dado à documentação de seus trabalhos um enfoque restrito aos seus interesses [Ping89]. A evolução e a diversificação dos sistemas computacionais também têm estimulado a produção de um sem número de documentos específicos, tanto no que concerne aos processadores de informação propriamente ditos, quanto aos dispositivos de apresentação e manipulação dessa informação. Löwgren [Löwg92] ressaltou esse aspecto fracionado das informações relativas ao projetos de interfaces e até mesmo propôs como alternativa o desenvolvimento de um *sistema gerenciador de interface de usuário (UIMS - User Interface Management System)* com uma base de conhecimento contendo informações relativas a fatores humanos.

A somatória desses fatores só reforça o fato de que a dispersão de informações relativas a aspectos particulares da interação homem-máquina dificulta a tarefa de confrontação de idéias complementares apresentadas em diferentes documentos, tornando o processo de seleção e consulta do projetista de interfaces cansativo e frustrante. Sem falar que o acesso simultâneo à diversidade de referências bibliográficas produzidas quase que diariamente por equipes de pesquisadores espalhadas pelo mundo inteiro não constitui uma das tarefas mais fáceis nem mais agradáveis. Daí a importância de se condensar em um único documento o maior número possível de diretrizes para projeto e avaliação de interfaces homem-máquina, facilitando o trabalho de consulta a tópicos genéricos e específicos e de atualização e complementação periódicas.

Teorias ou modelos, bem como princípios associados às diversas facetas da interação homem-máquina, oferecem condições para a organização do projeto como um todo [Alle82, Shne84, Cons99]. Diretrizes práticas - genéricas ou específicas, bem como regras de projeto e recomendações técnicas, embora difíceis de ser elaboradas (por conta da diversidade de opiniões e da subjetividade do contexto), representam uma contribuição importante de experiências positivas anteriores, um ponto de reflexão para o sucesso de novos trabalhos [Shne82, Sond82, Gain83, Shne84, Shne93], e referência para o desenvolvimento e a avaliação de novos produtos [Tetz91, Cons99]. Estratégias de teste e avaliação são, por outro lado, instrumentos que estimulam a pesquisa e conduzem o projetista à correção de decisões de projeto insatisfatórias, com base no grau de satisfação dos usuários finais do produto [Shne87, Este90].

Apesar de todos os benefícios que podem advir do uso de *diretrizes de projeto e guias de estilo* como instrumentos auxiliares em processos de desenvolvimento e avaliação de interfaces de usuário, a consideração de fatores humanos em projetos e avaliações de interfaces ainda é vista como um ingrediente crítico, assim como o é o próprio usuário [Mora81b, Cons99, Rask00]; além do mais, *quando*, *quanto* e *como* tal ingrediente deve ser incorporado a outros elementos de

projeto e avaliação ainda será encarado por muito tempo como um tópico polêmico, graças ao grau de subjetividade envolvido no contexto.

Outro fato que deve ser considerado é que embora tenham sido escritos muitos livros sobre tópicos de interfaces, poucos deles tratam de como especificá-las e/ou implementá-las [Rett92]. Root e Uyeda [Root93] traduziram em seu relatório sobre o *CHI'92 Styleguide SIG* a necessidade urgente, constatada pelos redatores de coletâneas de diretrizes e guias de estilo, de respaldo moral aos novos redatores, sob a forma de (i) recomendações sobre como elaborar guias de estilo e (ii) sugestões de consultas a bases científicas idôneas de informações que possibilitem a formulação de diretrizes confiáveis e consistentes.

Segundo estes autores, os desafios mais imediatos à comunidade pesquisadora no domínio de fatores humanos são: (i) *estudos mais consistentes e constantes de embasamento para a redação de diretrizes*; (ii) *mecanismos mais eficientes de transferência e difusão de resultados de pesquisa*, para que atinjam expressividade mundial; e (iii) *a evolução do desenvolvimento de guias de estilo e coletâneas de diretrizes a partir de metodologias sistemáticas*, que possibilitem a produção de boas recomendações e, por conseguinte, de boas interfaces.

Com relação a esses desafios, alguns autores [Smit86a, Cout88, Scap88, Poll90, Root92, Cons99] têm apontado sérios problemas de utilização das coletâneas de diretrizes e de guias de estilo como ferramentas de projeto e avaliação de interfaces de usuário e feito numerosas críticas ao tema, dentre as quais: (i) *falta de tempo do projetista ou do avaliador de interfaces para a leitura de documentos volumosos*; (ii) *generalidade excessiva* (ausência de indicações que auxiliem os projetistas/avaliadores na tradução de regras específicas pertinentes para o contexto do sistema de seu interesse) *ou especificidade excessiva* (recomendações relativas exclusivamente a casos particulares) *das recomendações contidas nos documentos*; (iii) *falta de precisão e clareza da importância relativa a ser dada às recomendações, se não de veracidade e consistência de seu conteúdo*; (iv) *implicação de interações entre critérios ergonômicos na aplicação de conjuntos de diretrizes correlatas*; e (v) *carência de revisão e atualização*. Tais problemas, separada ou conjuntamente, contribuem para a dificuldade de apreciação e aplicação destes documento por não especialistas.

Sneeringer [Snee78] comentou que a adoção de padrões familiares atua no sentido de atenuar a sensação de deslocamento e insegurança dos usuários sem experiência computacional prévia, além de representar uma transição mais "natural" da tarefa para o ambiente computacional. Por outro lado, Constantine e Lockwood [Cons99] reforçam esta questão ao afirmarem que bons padrões de interfaces de usuário e guias de estilo úteis emergem de requisitos do usuário tanto quanto da experiência adquirida pelos projetistas com a implementação desses requisitos. Ambas as considerações só vêm reforçar a necessidade de emprego de padrões de projeto que reflitam o melhor possível, explícita ou implicitamente, o uso cotidiano de recursos compartilhados pela comunidade de usuários à qual o projeto se destina, devendo as considerações duvidosas (adotadas) serem testadas em fases ligeiramente subseqüentes àquelas de prototipagem das diferentes camadas do produto, a fim de que se possa determinar o nível de satisfação das expectativas mais significantes da população-alvo de usuários.

Ohlson [Ohls78] e Foley [Fole84] ressaltaram a importância da seleção, pelo projetista da interface, dos dispositivos e técnicas através dos quais o usuário realizará tarefas de interesse, dentre os inúmeros subconjuntos que podem ser empregados para gerar bons projetos de interface [Bail93]. Diversos autores [Shne87, Niel90, Jeff91, Kara92, Bail93, Mand97, Cons99] concordam que um projeto satisfatório deve resultar de uma combinação de várias técnicas, restando apenas conhecer a extensão e as limitações de cada uma delas e aplicá-las segundo as conveniências.

Roberts [Robe66] já expressava em seu artigo a preocupação com a flexibilidade de um sistema computacional no tocante às aplicações e aos modos de interação. Compartilhando essa preocupação, Bass [Bass81] descreveu a concepção de uma interface generalizada para programas de aplicações, cuja estruturação denotava a preocupação com o uso de valores *default* e com o nível de auxílio aos usuários do produto. Em outro artigo [Bass85], o autor destacou a relevância do processo de "customização" da interface de usuário para aplicações específicas, bem como do desenvolvimento do projeto com base em um grau de flexibilidade suficiente para corresponder às alterações dos requisitos que ocorressem em fases posteriores e do gerenciamento das aplicações por profissionais intimamente familiarizados com o seu contexto (não necessariamente programadores ou projetistas de sistemas). Passada uma década e meia, constata-se que a facilidade de customização tem passado a ser um dos aspectos mais atentamente observados em processos de inspeção de interfaces de usuário.

Henninger *et al.* [Henn95] propuseram um método no qual organizações de desenvolvimento de *software* podem elaborar e refinar diretrizes específicas ao domínio com base nos tipos de aplicações que desenvolvem. De acordo com a exposição dos autores, o método facilita o processo de determinação do *quando* e *como* as diretrizes devem ser aplicadas a partir da restrição das diretrizes a casos específicos de projeto e fornecendo meios de compatibilização das exigências do usuário a técnicas específicas de interfaceamento que tenham se mostrado efetivas para usuários e domínios de aplicação similares. Partindo do argumento que casos concretos auxiliam os projetistas a interpretar diretrizes de projeto, tornando-as mais fáceis de compreender e aplicar a problemas de projeto de interesse, Henninger *et al.* [Henn95] validaram a metodologia proposta a partir de um sistema baseado em casos, denominado *Mimir*, usado como suporte ao refinamento e enquadramento de diretrizes a casos relevantes.

Um recurso que tem se popularizado no âmbito das atividades de projeto e desenvolvimento de interfaces gráficas de usuário tem sido as ferramentas destinadas à implementação de interfaces (*interface builders*). Puerta *et al.* [b] argumentam que as ferramentas típicas para a prototipagem e implementação de interfaces são centradas em engenharia e atuam principalmente em janelas e *widgets*, não oferecendo um suporte específico para projetos de interfaces centrados no usuário. Os autores desenvolveram um editor de *layouts* de interfaces baseado em modelos, o MOBILE (*Model-Based Interface Layout Editor*), uma ferramenta de prototipagem de interfaces que dá suporte a projetos centrados no usuário, orientando o processo de implementação da interface a partir de modelos usuário-tarefa e de uma base de conhecimento de diretrizes de projeto de interfaces.

Diversos autores⁵² expressaram, segundo diferentes enfoques, a importância da consideração de regras, princípios heurísticos e/ou técnicos, recomendações e/ou diretrizes de projeto e de avaliação, relacionadas tanto com aspectos operacionais quanto com fatores humanos, i.e., a necessidade de um respaldo bibliográfico para o desenvolvimento e avaliação de processos e/ou produtos de *software*.

Ao longo das duas últimas décadas, um contingente significativo de pesquisadores⁵³ tem se preocupado com a compilação de diretrizes genéricas e específicas de projeto, quer com a revisão e atualização de trabalhos envolvendo o uso de diretrizes de projeto, quer com o desenvolvimento de ferramentas computacionais, fundamentadas em diretrizes de projeto, destinadas ao desenvolvimento de interfaces homem-máquina.

O crescimento exponencial e a difusão capilar da Web vem proporcionando a emergência de uma nova geração de aplicações [Niel00], caracterizada por uma relação direta entre negócios e consumidores, cuja implementação híbrida o desenvolvimento de sistemas de informação tradicionais com aplicações de hipermídia [Frat99]. Esta geração de aplicações tem representado um desafio para os projetistas de ferramentas de *software*, assim como para os pesquisadores de abordagens típicas destinadas à produção de *software*, atraindo atenção e capital como nenhuma outra tecnologia emergente nas últimas décadas, e dividindo a opinião dos especialistas [Shne98].

Neste contexto florescente e promissor, em que a complexidade das interfaces se alia à diversidade de usuários e tarefas, o desenvolvimento centrado no usuário ocupa cada vez mais um papel de primeira grandeza e os procedimentos avaliatórios de projetos em andamento se fundamentam cada vez mais nas exigências do usuário. A definição dos objetivos empresariais, do contexto de uso almejado e dos cenários-chaves de uso, ocupam o primeiro plano do processo de desenvolvimento. A estrutura dos *sites* e o projeto das páginas são submetidos à avaliação por usuários finais representativos. A gestão e a manutenção das páginas figuram na lista de aspectos a serem cuidadosamente observados na manutenção da usabilidade dos *sites*. Segundo Bevan [Beva98], em um cenário desta natureza os projetos carecem imprescindivelmente do suporte de diretrizes direcionadas para o estilo de redação, a navegação e o projeto de páginas adotados correntemente pela Web.

Morkes e Nielsen [Mork98] enfatizaram aspectos relativos à concisão, à facilidade de exploração e a objetividade do estilo de redação para a Web. Os autores discutiram questões de reprojeção de conteúdo da Web, tendo afirmado que um estudo comparativo do *site* original com o *site* reescrito atestou a otimização da usabilidade deste último em 159% em relação ao primeiro.

⁵² E.g. Kroll [Kro171], Shneiderman [Shne79, Shne82, Shne84, Shne87, Shne00], Galitz [Gali85], Grudin [Grud89, Grud90, Grud91], Nielsen [Niel90, Niel92b, Niel93c, Niel00], Mitchell e Shneiderman [Mitt89], Dzida [Dzid89], Nickerson e Pew [Nick90], Jeffries *et al.* [Jeff91], Karat [Kara92], Mayhew [Mayh92, Mayh99], Gill93, Gray [Gray93], Treu [Treu94], Mandel [Mand97], Constantine e Lockwood [Cons99].

⁵³ E.g. Shneiderman [Shne87], Galitz [Gali85], Grudin [Grud89, Grud90, Grud91], Smith e Mosier [Smit86b], Dzida [Dzid89], Mitchell e Shneiderman [Mitt89], Norcio [Norc89], Tognazzini [Togn90], Jeffries *et al.* [Jeff91], Tetzlaff e Schwartz [Tetz91], Mayhew [Mayh92], Gray [Gray93], Giliand [Gill93], Queiroz [Quei94], Cohen *et al.* [Cohe95b], Dilli & Hoffmann [Dill95], Gorny [Gorn95], Ogawa e Ueno [Ogaw95], Crow [Crow95], Iannella [Iann95], Reiterer [Reit95], Cohen [Cohe95a], Vanderdonck [Vand95], Constantine e Lockwood [Cons99] e Nielsen [Niel00].

Além do mais, uma sondagem revelou que a satisfação subjetiva dos usuários do *site* reprojeto foi superior àquela dos usuários do *site* original, bem como que o desempenho usando o *site* reescrito foi melhor, em termos do tempo de execução da tarefa, número de erros cometidos e carga cognitiva.

Em um levantamento realizado sobre projetos de páginas para a Web, Vora [Vora98] constatou que quase 90% do respondentes adotavam diretrizes de projeto e/ou guias de estilo ao projetarem páginas para a Web: cerca de 38% usavam guias de estilo de projeto desenvolvidos na empresa e aproximadamente 50% deles usavam tanto diretrizes de projeto disponíveis na Web quanto diretrizes de projeto elaboradas na empresa. Segundo Vora, conquanto tal fato seja encorajador, só permite imaginar até que ponto as diretrizes de projeto são usadas em projetos de páginas da Web, a julgar por uma pesquisa semelhante conduzida por Mosier e Smith [Mosi86] em nível de projetos de aplicações de *software*, a qual mostrou que as diretrizes de projeto de interfaces eram, de um modo geral, consideradas úteis, apesar de apresentarem problemas significativos no tocante à aplicação prática.

Dix [Dix99] examinou várias facetas do projeto de interfaces para a Web em nível mundial, considerando aspectos como (i) o uso da Web como uma plataforma para prototipagem ou uso efetivo de interfaces usuário-computador; (ii) diretrizes e recomendações destinadas ao projeto de sites da Web; (iii) a natureza da Web enquanto mídia; (iv) o modo através do qual as interfaces podem auxiliar os usuários a administrarem a complexidade da Web, inclusive mecanismos de históricos e revisões diagramáticas. A partir deste último tópico, Dix conduziu uma discussão das geometrias múltiplas da Web, constituídas por links, conteúdos e atividades de navegação por usuários. Dix concluiu com uma reflexão sobre o futuro da Internet como um fenômeno de integração, incluindo uma descrição de *onCue* da *aQtive*, uma aplicação que integra serviços da Internet, aplicações desktop e atividades correntes de usuário, e prevendo a longo prazo a operacionalização da *PopuNET*, uma rede disponível a todos, em toda parte e o tempo todo.

Fraternali [Frat99] realizou uma revisão do estado da arte das ferramentas de desenvolvimento de aplicações para a Web, tanto em nível comercial quanto no campo da pesquisa, identificando e caracterizando diferentes categorias de soluções, avaliando sua adequação às exigências de desenvolvimento de aplicações para a Web e esclarecendo questões relativas ao processo, aos modelos, linguagens e notação, ao reuso, à arquitetura e à usabilidade, finalizando com a apresentação de possíveis tendências futuras.

Por sua vez, Nielsen [Niel00] apresenta várias diretrizes de projeto de interfaces para a Web, assim distribuídas: (i) *títulos de links* (Chapter 2, pp. 60-62); (ii) *folhas de estilo* (Chapter 2, pp. 84-85); (iii) *redação para a Web* (Chapter 3, pp. 101-115); (iv) *títulos* (Chapter 3, pp. 124-125); (v) *legibilidade* (Chapter 3, pp. 125-126); (vi) *redação de documentação online* (Chapter 3, p. 131); (vii) *escopo de buscas* (Chapter 4, p. 225); e (viii) *elaboração de padrões de projeto de interfaces intranet* (Chapter 4, pp. 281-284).

Dentre os autores que relataram recentemente pesquisas envolvendo o uso de diretrizes de projeto estão Jamieson e Vicente [Jami98] e Lehane et al. [Leha00].

Tanto os *guias de estilo* quanto os *padrões de projeto* têm se tornado objeto de atividades intensivas nos últimos anos [Hold89, Aber90], tendo em vista que, sendo o

fundamento de cada avanço tecnológico, os padrões técnicos estabelecem o elo de continuidade entre cada inovação e as precedentes através da referência que fazem a outros padrões técnicos anteriores [Krec96]. Inovações tecnológicas consolidadas impulsionam, em menor ou maior grau, o fluxo do progresso: enquanto a maior parte provoca apenas leves ondulações no fluxo, algumas formam ondas que desencadeiam profundas alterações tecnológicas. Os padrões técnicos representam um meio de delineamento dessas ondas [Krec96], tornando consistentes os avanços da tecnologia.

Em termos de popularidade, o uso de guias de estilo e padrões bem concebidos ocupa, segundo Constantine e Lockwood [Cons99], o segundo lugar na lista de estratégias adotadas para a otimização da usabilidade, seguindo de perto os ensaios de usabilidade. Tal fato se dá sobretudo porque guias de estilo e padrões bem concebidos podem conferir (i) consistência ao processo de desenvolvimento de interfaces, possibilitando a geração de produtos mais fáceis de aprender e manipular; (ii) melhores práticas de projeto e avaliação aos profissionais envolvidos com o desenvolvimento de interfaces usuário-computador, promovendo a sistematização mais efetiva de ambos os processos; e (iii) eficácia e eficiência à gestão de recursos financeiros, físicos, materiais e humanos envolvidos no contexto mais abrangente de desenvolvimento de produtos, não apenas de suas interfaces com o usuário.

Gale [Gale96] ressaltou o uso de *guias de estilo* apropriados como ferramentas vitais na otimização de interfaces gráficas de usuário (*GUI*), descrevendo uma iniciativa conjunta para o seu desenvolvimento por projetistas de interfaces gráficas com a colaboração de usuários finais. Neste documento, o autor também apresentou uma revisão geral de guias de estilos, destacando os benefícios associados ao seu uso em projetos de interfaces, as principais razões que contribuem para a ineficácia dos guias de estilo como instrumentos auxiliares de projeto, e as estratégias que podem ser adotadas para operacionalizá-los em contextos de desenvolvimento de interfaces gráficas.

Ao final da década de 80, Holdway e Bevan [Hold89] já discutiam a importância dos padrões internacionais, destacando aspectos como utilidade, dificuldade de interpretação, atualização e evolução e descreviam as necessidades de padronização em nível internacional, com fins à consistência, à otimização da usabilidade, ao conforto e bem-estar do usuário e à inspeção de produtos. Os autores citavam as atividades recentes de padronização, descrevendo as organizações envolvidas - ISO, DIN e ANSI.

Meia década depois, Bevan e Holdway [Beva93] expressaram a necessidade de padronização, do ponto de vista dos usuários de sistemas interativos. Adicionalmente, Bevan [Beva95c] discutiu o caráter genérico dos resultados das iniciativas de padronização internacional dos processos interativos usuário-computador, partindo da argumentação de que, embora oferecessem o benefício da consistência, as interfaces de usuário padronizadas tornavam-se obsoletas devido aos avanços tecnológicos. Neste artigo, Bevan associou as iniciativas de padronização às abordagens *top-down* e *bottom-up* de desenvolvimento de *software* (discutidas no capítulo anterior), argumentando que os padrões concebidos sob a visão global da qualidade (abordagem *top-down*) são aplicáveis ao contexto abrangente de projeto e objetivos de qualidade, enquanto os padrões elaborados segundo a visão orientada ao produto (abordagem *bottom-up*) se prestam melhor ao projeto de atributos específicos, sendo mais relacionados às necessidades do

projetista de interfaces e à importância da usabilidade para o desenvolvimento de produtos de *software*. O autor também comentou os padrões ISO/IEC 9126, 11581 e 14598, assim como os padrões ISO 8402, 9000 e 9241, evidenciando os conceitos de *qualidade global* e *qualidade de uso* (também discutidos no capítulo anterior).

Destaca-se no âmbito do desenvolvimento de padrões de interfaces os esforços e iniciativas da *International Standards Organization* - ISO [Broo90, Niel93b] e da Comunidade Européia [Stew90], além dos trabalhos de diversas organizações para a padronização em nível nacional [Dzid89, Dzid96], e.g. *ANSI* (EUA), *DIN* (Alemanha), *ABNT* (Brasil), *BSI* (UK). No tocante ao desenvolvimento interfaces padronizadas, a comunidade usuária vem se familiarizando e usando os resultados das iniciativas de diversas organizações de desenvolvimento de produtos de *software* [Berr88, Niel89a, Good89, Togn89], cujas interfaces refletem a padronização segundo estilos formalmente documentados (e.g., *SAACUA* da *IBM* [IBM91a, IBM91b], *WINDOWS* da *Microsoft* [Micr95], *MAC OS 8* da *Apple* [Appl92, Appl97], *MOTIF* da *OSF* [OSF90], *JAVA* da *Sun Microsystems Inc.* [Sun99]). Tais iniciativas denotam o interesse que vem se generalizando por *padrões* de processamento da informação [Berg90], embora também se possa perceber que parte deste interesse se deve à evidência da importância de considerações de usabilidade no desenvolvimento e avaliação de projetos de interfaces de usuário, sobretudo no que diz respeito ao aspecto *consistência* [Niel89a, Togn90].

Segundo Nielsen [Niel93b], *padrões* de interfaces são importantes para a avaliação da consistência de uma interface, porque além de suas características gerais, possibilitam às equipes de projeto e avaliação de produtos de *software* um suporte à elaboração de seus próprios padrões *ad hoc*. Nielsen [Niel92a, Niel93b] ainda ressalta a distinção existente entre *padrões* e *diretrizes de projeto* com o argumento de que *padrões* especificam como as interfaces deveriam ser apresentadas ao usuário, enquanto *diretrizes de projeto* apenas informam sobre as características de usabilidade que tais interfaces deveriam apresentar, listando princípios bem conhecidos a serem considerados em atividades de desenvolvimento e avaliação de projetos de interfaces usuário-computador. Em ambos os casos, no entanto, o projetista deverá julgar a importância relativa do conteúdo informativo disponível, de acordo com as circunstâncias da tarefa e com aspectos organizacionais da situação a ser solucionada [Dzid89].

Peddie [Pedd92] afirmou que os *padrões* para linguagens de programação e ferramentas gráficas apresentam muitos benefícios, e.g., a visualização tridimensional de dados, a difusão de ferramentas de *software* "abertas", embora também hajam inconvenientes e riscos, dentre os quais pode ser citada a implementação inexpressiva e deficiente de um padrão. Além do mais, embora os *padrões* de interface estejam intimamente ligados a questões relativas à facilidade de aprendizado e de uso, o que reduz custos com cursos de treinamento [Pols88], eles são difíceis de aplicar, tanto em tarefas de projeto, quanto em atividades de avaliação, de modo que uma das linhas de pesquisa para solução deste inconveniente aponta para a avaliação automática da conformidade de projetos com um padrão [Löwg92]. Entretanto, ainda há muitos aspectos de usabilidade que não se prestam à representação e análise automatizada [Niel93c].

Em um nível mais sinóptico, Bach [Bach95] discriminou três domínios para os quais a comunidade americana envolvida com iniciativas de padronização tem focalizado a atenção - a Internet, a indústria tradicional e as telecomunicações, delineando as fraquezas e as forças de cada domínio nesse sentido. Segundo Bach, a comunidade voltada para a padronização da

Internet tem criado seu próprio processo, buscando tirar proveito tanto das tecnologias usadas pelos pioneiros da Internet quanto das filosofias por estes advogadas (ver também Crocker [Croc93]). Por outro lado, os elaboradores de padrões tradicionais se espelham nas iniciativas de padronização do ANSI e nos esforços deste instituto no sentido de melhorar o ambiente empresarial constantemente em transição. Por fim, Bach comenta os esforços dispendidos pelas organizações de telecomunicações, que simbolizam os esforços da comunidade de tecnologia da informação, no sentido de desenvolver padrões que atendam de modo integrado as necessidades tradicionais e as tecnologias emergentes. O que se apreende destas iniciativas é que há um espaço promissor em qualquer um dos três domínios para as iniciativas de desenvolvimento de padrões para o projeto e a avaliação de interfaces usuário-computador, voltados para a usabilidade.

No âmbito das iniciativas de padronização de interfaces usuário-computador, Dzida [Dzid95, Dzid96] enfatizou a diferença entre o desenvolvimento de interfaces padronizadas e o desenvolvimento de padrões para interfaces. As iniciativas de desenvolvimento de interfaces padronizadas, como já foi mencionado, resultam em guias de estilo, documentos mais específicos aos "atrativos visuais e impressivos" da interface - o *look and feel* - através dos quais o usuário estabelece a comunicação com o sistema. Por outro lado, as iniciativas de desenvolvimento de padrões para interfaces resultam em documentos mais genéricos, de mais alto nível, através dos quais se pode desenvolver interfaces usáveis, independentemente do estilo a ser considerado, ou avaliar a usabilidade de interfaces desenvolvidas de acordo com diferentes estilos. Rosenzweig [Rose96] descreveu o processo conduzido em 1994 pela *Eastman Kodak Company*, com o propósito de desenvolver diretrizes e padrões de projeto corporativos para seus produtos de *software* para computadores pessoais. No tocante a iniciativas desta natureza, Nielsen [Niel96a] chamou atenção para a necessidade de consideração da complexidade do comportamento humano, em especial a faceta que reflete a dependência contextual, em processos de elaboração de diretrizes e padrões corporativos.

No tocante aos padrões internacionais, estes são, além de genéricos, naturalmente descritivos, haja visto refletirem opiniões e experiências de um grande número de especialistas espalhados pelo mundo inteiro. Adicionalmente, os padrões internacionais apresentam um caráter prescritivo, uma vez que se pretende que as opiniões e experiências neles registradas sejam adotadas pelo maior número possível de profissionais daquela área de atuação [Rehe96].

Quanto mais madura e consolidada for a área empreendedora dos esforços de padronização, mais o padrão elaborado refletirá resultados de práticas comuns naquela área. Por sua vez, a adoção de um padrão internacional, exceto em casos de decreto, só ocorrerá se o padrão for realmente consistente e atender as necessidades do profissional típico da área. Rehesaar [Rehe96] descreveu detalhadamente o processo de desenvolvimento de padrões internacionais para a engenharia de *software* pelo *JTC1*, um comitê formado em 1987 pela ISO e IEC, ressaltando o rigor dos processos de elaboração e a composição internacional de seus sub-comitês. O autor também observou o rigor e a atenção dados por este comitê à apresentação de soluções práticas para os problemas inerentes à engenharia de *software*, uma área relativamente recente e, em alguns aspectos, ainda imatura. Vale a pena salientar que este mesmo comitê, juntamente com os comitês 159 e 184, é responsável pelo desenvolvimento de padrões internacionais para a engenharia da usabilidade.

Em nível da aplicação de padrões internacionais, Douglas et al. [Doug99] relataram a aplicação do *draft* da Parte 9 do padrão ISO 9241, que normatiza o teste de dispositivos de apontamento, a fim de avaliarem a validade científica e a praticidade das dimensões recomendadas pelo padrão – *desempenho* e *conforto*. Foram utilizados dois dispositivos de apontamento para *laptops*, um *joystick* isométrico e um *painel sensível ao toque* (*touchpad*). Os autores apresentaram, além da descrição da metodologia adotada e dos resultados obtidos, uma discussão dos problemas encontrados na adoção da Parte 9 do padrão ISO 9241 e recomendações destinadas à revisão e otimização da referida parte.

3.2.2.3 Avaliação Heurística (*Heuristic Evaluation*)

Segundo Jeffries et al. [Jeff91], a *avaliação heurística* (*heuristic evaluation*) consiste do estudo aprofundado de um produto, conduzido por especialistas, com base em experiência e conhecimentos pessoais, com o propósito de identificar propriedades que possam traduzir problemas na usabilidade de tal produto. Tal ponto de vista é corroborado por Treu [Treu94], que lhe acrescenta a necessidade de inclusão de uma heurística escrita para guiar as atividades da equipe de avaliadores.

Nielsen [Niel90, Niel92a, Niel92b, Niel93b, Niel93c, Niel94a, Niel95] visualiza o processo de um modo um pouco menos ortodoxo. Embora concorde com a necessidade de alguma experiência dos avaliadores no que diz respeito aos princípios que serão aplicados nas mais diferentes circunstâncias durante a descoberta de problemas de usabilidade, Nielsen também ressalta que mesmo não especialistas podem detectar diversos "gargalos" na usabilidade de um produto a partir da *avaliação heurística*. Ao invés de um grupo de especialistas, ele recomenda a prática desta estratégia avaliatória por vários indivíduos diferentes, visto que diferentes competências podem detectar categorias diferentes de problemas de usabilidade.

Jeffries et al. [Jeff91] comentaram em seu estudo comparativo a divergência da *análise heurística* por eles concebida e a abordagem adotada por Nielsen, no que diz respeito à prática da *avaliação heurística* por projetistas de interfaces. Neste estudo comparativo, foram convidados 04 avaliadores, membros de um grupo de pesquisa em interação usuário-computador que, segundo a experiência dos autores, apresentavam perfis cognitivos e competências profissionais⁵⁴ mais adequadas ao contexto do que membros de uma equipe de projeto de interfaces. Após um período de 02 semanas, os avaliadores concluíram suas atividades, cada um dos quais tendo realizado sessões de avaliação em intervalos de tempo pré-definidos. Estes intervalos foram registrados em seus respectivos relatórios finais. Infelizmente, os autores não comentaram em seu artigo as heurísticas adotadas pelos avaliadores.

Concebida com propósitos de análise, diagnóstico e, quando possível, prognóstico do produto avaliado, a *avaliação heurística* se respalda em listas de regras e princípios de usabilidade, denominados *heurísticas de usabilidade*, essencialmente similares às recomendações e/ou normas contidas, respectivamente, em documentos típicos de diretrizes de projeto e padrões [Niel90, Niel92a, Niel93b]. No entanto, estes documentos costumam ser extensos e focalizar um grande número de recomendações e/ou regras a serem observadas, o que os torna muitas vezes

⁵⁴ Foram convidados profissionais especializados em ciência comportamental e engenharia da usabilidade, com anos de experiência na prática de *avaliação heurística* de interfaces de usuário.

inadequados para uso em ensaios heurísticos. De um certo modo, a intuição associada ao senso comum e ao conhecimento armazenado na memória de longa duração possibilita à maioria dos indivíduos emitir opiniões sobre o que é adequado ou não em um produto, dentro de contextos específicos. Tais iniciativas constituem bases para algum tipo de avaliação heurística.

A fronteira que separa o julgamento respaldado em padrões e diretrizes daquele fundamentado em heurísticas de usabilidade é bastante tênue. Tão tênue, na verdade, que alguns especialistas (e.g., Nielsen [Niel92a, Niel93b]) consideram os princípios listados nas coletâneas de diretrizes uma espécie de alicerce para a *avaliação heurística*. Talvez a sutil diferença entre estas estratégias avaliatórias resida no fato de que muitos dos princípios bem conhecidos que preenchem páginas e páginas de documentos de diretrizes já estejam armazenados na memória de longa duração dos especialistas em heurísticas, permitindo-lhes julgar um determinado produto sem precisar consultar uma coletânea escrita. Por outro lado, talvez resida *apenas* na etapa considerada de concepção do produto: as coletâneas de diretrizes são mais empregadas nas etapas de concepção e implementação, enquanto as heurísticas, que sintetizam a essência de tais coletâneas, se prestam melhor na etapa de inspeção da implementação.

O processo de *avaliação heurística* se inicia com a inspeção individual da interface pela equipe avaliadora. Os resultados individuais só são confrontados, discutidos e integrados após todas as atividades de avaliação terem sido concluídas, a fim de garantir a independência e a não polarização das opiniões de cada especialista. O registro do processo avaliatório pode se dar através de mecanismos audiovisuais ou seguir os moldes de apresentação dos relatórios convencionais. Relatórios escritos requerem esforços adicionais dos avaliadores, além da presença de um supervisor do processo, cujo encargo é a leitura dos relatórios individuais e a redação de um documento final contendo todos os aspectos avaliados concatenados num texto fluente e homogêneo.

Uma alternativa ao método acima é a inclusão de um observador durante as sessões de avaliação, enquanto cada avaliador verbaliza suas opiniões e descobertas, à medida que estuda o produto. Tal opção introduz como inconveniente a necessidade de mais um componente, embora a vantagem de reduzir a carga de trabalho dos avaliadores e de se poder contar com os resultados finais imediatamente após o processo haver sido concluído possa representar benefícios mais do que compensadores. Se o observador escolhido estiver familiarizado com o aplicativo em teste, sua assistência em casos de problemas em sua manipulação ou mesmo na explanação de determinados aspectos de sua interface também poderá ser valiosa.

Durante um ensaio de usabilidade típico, o observador necessita interpretar as ações do usuário, a fim de inferir como se relacionam com os problemas de usabilidade avaliados, enquanto numa abordagem heurística seu papel se restringe ao registro dos comentários de cada avaliador sobre a interface avaliada e à organização de todas as suas anotações em um único documento. Além do mais, num ensaio de usabilidade o observador visa a detecção de falhas cometidas pelo usuário durante o uso do aplicativo avaliado, o que os faz redobrar a cautela ante o esclarecimento de dúvidas do usuário ou à prestação de assistência quanto a determinados problemas surgidos durante a observação do usuário. No processo de *avaliação heurística*, pelo contrário, esclarecimentos e assistência aos avaliadores são imprescindíveis, pois agilizam o processo avaliatório e o complementam.

Atividades típicas de *avaliação heurística* envolvem o exame minucioso da interface, repetido várias vezes, a inspeção dos mais diversos aspectos do diálogo e a comparação das características da interface com uma lista de princípios de usabilidade previamente elaborada para o contexto ou já validada em outros processos avaliatórios. Tais listas são compostas de regras gerais que tentam descrever propriedades comuns de interfaces usáveis. Às heurísticas genéricas de verificação de todos os elementos de diálogo, o avaliador também pode incluir princípios de usabilidade adicionais ou ainda resultados, conhecimentos e experiências que tenha em mente e que possam lhe ser úteis na análise de características específicas pertinentes ao contexto.

Nielsen e Molich [Niel89b, Niel90] e Nielsen [Niel92a, Niel93b, Niel93c] apresentaram em seus trabalhos a lista de princípios de usabilidade por eles adotada em processos de *avaliação heurística*, a qual recomendam a todos os projetistas de interfaces usuário-computador. Por outro lado, encontra-se na literatura da área autores que descreveram experimentos envolvendo *avaliação heurística*, embora não mencionem as heurísticas adotadas, e.g. Jeffries *et al.* [Jeff91], Desurvire *et al.* [Desu91, Desu92]).

Nielsen [Niel91] comentou que a usabilidade não podia ser definida sem o conhecimento do contexto em que o sistema projetado seria usado, ressaltando a inexistência de critérios *absolutos* de usabilidade. Sendo este um conceito genérico que não pode ser mensurado, apenas avaliado a partir de diversos critérios mensuráveis, Nielsen [Niel94a, Niel94b] associou a usabilidade a parâmetros que se enquadram em duas grandes categorias, a saber: *medidas da preferência subjetiva do usuário*, que denotam a preferência do usuário pelo sistema; e *medidas de desempenho objetivo*, que estimam quão capazes os usuários são de interagir com o sistema.

Nielsen [Niel94c] comparou várias listas publicadas de heurísticas de usabilidade com um banco de dados de problemas de usabilidade, registrados a partir de diversos projetos previamente avaliados, a fim de determinar quais heurísticas explicavam melhor os problemas de usabilidade correntes. Foram utilizados dados relativos a **249** problemas, anteriormente identificados a partir de avaliações conduzidas no âmbito de **11** projetos: (i) **7** inspecionados a partir de heurísticas e **4** a partir de ensaios envolvendo usuários de teste; (ii) **4** avaliados em uma fase inicial do ciclo de vida de desenvolvimento do produto e **7** avaliados em uma fase avançada do desenvolvimento; e (iii) **2** com interfaces textuais, **6** com interfaces gráficas e **3** com interfaces operadas por telefone. Cada uma das **101** heurísticas de usabilidade consideradas foi avaliada através de uma escala de **5** pontos⁵⁵ conforme seu potencial explicativo de cada um dos problemas registrados no banco de dados. A partir de uma análise fatorial das explicações e da capacidade da heurística de cobrir o mais extensamente possível os problemas analisados, Nielsen propôs uma nova lista de **9** heurísticas: *visibilidade do estado do sistema*, *compatibilidade entre o sistema e o mundo real*, *controle e liberdade do usuário*, *consistência e padronização*, *prevenção de erros*, *reconhecimento em vez de anulação de ações*, *flexibilidade e eficiência de uso*, *projeto estético e minimalista* e *ajuda ao usuário no reconhecimento, diagnóstico e recuperação de erros*.

⁵⁵ 0 = não explica o problema

1 = se relaciona superficialmente com algum aspecto do problema

2 = explica uma pequena parte do problema, mas não seus principais aspectos

3 = explica os principais aspectos do problema, mas há alguns aspectos que não consegue explicar

4 = explicação quase completa do *porque* do problema, mas ainda há aspectos que transcendem a explicação

5 = explicação completa do *porque* do problema

Muller *et al.* [Mull95] descreveram uma extensão e validação da abordagem avaliatória heurística proposta por Nielsen e Molich [Niel90, Niel92c], à qual incluíram aspectos "humanísticos" de sistemas computacionais, sob a forma de três heurísticas adicionais relativas à qualidade do trabalho com o produto, qualidade da sessão de trabalho e respeito para com as habilidades do usuário. As referidas heurísticas foram incluídas e validadas em um processo participativo de avaliação heurística do LEAP (*Learn, Explore and Practice*), um sistema inteligente de tutoria, que contou com a participação de 5 especialistas em fatores humanos e 3 especialistas na área do produto. Dentre os problemas detectados pelos avaliadores, 33 % se fundamentaram apenas nas 10 heurísticas propostas por Nielsen *et al.* [Niel92c], 52 % se fundamentaram em pelo menos uma das heurísticas de Nielsen *et al.* [Niel92c] e em pelo menos uma das novas heurísticas propostas e 15 % se fundamentaram exclusivamente em uma ou mais das novas heurísticas propostas. No tocante às recomendações emitidas a partir da avaliação heurística, 31 % se fundamentaram apenas nas 10 heurísticas propostas por Nielsen *et al.* [Niel92c], 59 % se fundamentaram em pelo menos uma das heurísticas de Nielsen *et al.* [Niel92c] e em pelo menos uma das novas heurísticas propostas e 10 % se fundamentaram exclusivamente em uma ou mais das novas heurísticas propostas, atestando uma contribuição significativa das três heurísticas adicionais.

No contexto do projeto de páginas para a Web, Nielsen [Niel96b] discutiu 10 falhas consideradas, em sua opinião, como as mais relevantes e correntes. A discussão dessas falhas, relativas a questões diversas, tais como navegação, redação, tempo de *download*, atualização da informação, etc, foi posteriormente complementada em (Nielsen [Niel96c, Niel96d, Niel96e]) e atualizada em (Nielsen [Niel97d, Niel97e, Niel97f, Niel97g, Niel97h]).

Levi e Conrad [Levi96a, Levi96b] realizaram uma avaliação heurística do protótipo de um sistema de acesso público à Web contendo cerca de 100 páginas em formato HTML e 3 mini-aplicações embutidas, a saber: (i) formulário de sugestões e comentários; (ii) mecanismo de busca textual através de resumos de artigos de pesquisas econômicas e estatísticas; e (iii) interface baseada em formulários para o *Consumer Price Index*. A avaliação heurística foi conduzida por dois grupos de indivíduos, um deles constituído de 4 especialistas em interfaces usuário-computador e outro composto por 4 projetistas pertencentes à equipe de implementação do protótipo. Os autores apresentaram um conjunto de 8 heurísticas de usabilidade, empregadas no processo de avaliação do protótipo e derivadas das heurísticas propostas por Nielsen [Niel94a], de princípios específicos a hipertexto propostos por Shneiderman e Kearsley [Shne89, *in* Levi96b] e de considerações intuitivas dos avaliadores. Segundo os autores, o conjunto de heurísticas resultante da integração dos estudos de Nielsen [Niel94a] e Shneiderman e Kearsley [Shne89] com as considerações empíricas dos avaliadores, pode ser aplicado a contextos avaliatórios similares.

Gerhardt-Powals [Gehr96] também propôs um conjunto de dez princípios de projeto e avaliação de interfaces para a Web, direcionados para a qualidade e rapidez do processo de interação do usuário com os recursos oferecidos pelos *sites* da Web.

Doubleday *et al.* [Doub97] descreveram o modo como a interface *INTUITIVE (Interactive User Interface Tools In a Visual Environment)* do projeto *ESPRIT* suporta o esforço cognitivo envolvido em tarefas de recuperação de informações. No contexto da revisão de métodos de avaliação de uso corrente, os autores relataram séries de experimentos estruturados com o propósito de comparar métodos de teste da usabilidade aplicados na avaliação da interface

INTUITIVE. Doubleday *et al.* [Doub97] apontaram como propósitos-chaves da iniciativa (i) a investigação de problemas defrontados por usuários durante a execução de tarefas de recuperação de informações; e (ii) a avaliação de métodos de avaliação sob a perspectiva do foco dos problemas, a qualidade dos resultados e a efetividade de custo de cada método. Neste estudo, os autores confrontaram os resultados de ensaios de usabilidade envolvendo observação direta, ensaio retrospectivo e uso de questionários com resultados obtidos a partir de sessões de avaliação heurística fundamentadas nas heurísticas propostas por Nielsen [Niel93b, Niel94a, Niel97c].

Um aspecto importante a considerar em processos tanto de projeto quanto de avaliação é que, independentemente do conjunto de heurísticas que se adote, estas deverão refletir critérios de usabilidade reconhecidos e práticas vigentes. Tendo em vista que a engenharia da usabilidade ainda é uma área emergente e que tanto os usuários quanto as tecnologias de *hardware* e *software* estão em contínuo processo de evolução, heurísticas consideradas atualmente como apropriadas a determinados contextos poderão não ser apropriadas em um futuro próximo, embora possam ser revisadas, complementadas e atualizadas, a fim de conservarem sua aplicabilidade àqueles contextos.

O processo contínuo de refinamento de heurísticas de usabilidade pode implicar diferenças positivas e relevantes na usabilidade dos produtos nelas fundamentados. Infelizmente, como bem ressaltaram Catani e Biers [Cata98], freqüentemente os testes de desempenho conduzem à identificação de "problemas de usabilidade" substancialmente distintos daqueles explicados a partir de heurísticas de usabilidade. Estes "novos problemas" poderão ser, evidentemente, um reflexo da adoção de listas de heurísticas incompletas, desatualizadas ou falhas. Catani e Biers [Cata98], assim como Jacobsen *et al.* [Jaco98] e Zhang *et al.* [Zhan98b, Zhan98c], também propuseram heurísticas de usabilidade.

Toleman e Toleman [Tole98] relataram uma avaliação heurística modificada que fundamentou-se no projeto intitulado *Student Electronic Services and Support*, posteriormente denominado *USQconnect*. Os autores adaptaram a metodologia descrita por Nielsen [Niel94a], reduzindo o número de avaliadores, usualmente entre três a cinco, para um único avaliador. No processo de inspeção foram consideradas as dez heurísticas de usabilidade propostas por Nielsen [Niel93b, Niel94a, Niel97c] e a lista dos dez principais equívocos cometidos em projetos para a Web [Niel96b]. O mérito da adaptação reside no fato de haver resultado em uma estratégia econômica de avaliação da usabilidade (*discount usability evaluation*) que, segundo a discussão de Toleman e Toleman [Tole98], pode ser adotada em contextos nos quais a escassez de avaliadores e/ou de recursos econômicos é um fator limitante.

Uma área que vem crescendo em nível de importância é o comércio eletrônico (*e-commerce*) que por se fundamentar na Web e ser ainda emergente, carece de iniciativas de pesquisa de interfaces usuário-computador, tanto em nível de projeto quanto de avaliação. O comércio eletrônico baseia-se, *grasso modo*, em navegadores da Web (*Web browsers*), constituindo-se de *sites* criados para a transação de bens e serviços via Internet, embora os negócios também possam ocorrer via fax, telefone ou outros meios indicados no *site* consultado. Aplicações de comércio eletrônico vem surgindo a taxas vertiginosas, sobretudo nos últimos três anos, a ponto de organizações de pesquisa do mercado internacional, e.g., *IDC Market Research*,

ActiveMedia e a Forrester Research, estimarem que as transações comerciais mundiais excederão 1 trilhão de dólares em 2003 [Rohn98].

Motivada pelo potencial que esta área oferece para a pesquisa de usabilidade de interfaces usuário-computador, Rohn [Rohn98] realizou uma pesquisa do tema, que entre outros resultados apresentou uma lista de heurísticas, destinadas à produção de melhores interfaces para sites de comércio eletrônico, derivadas de diversos procedimentos avaliatórios e de recomendações de diversas publicações. As heurísticas propostas por Rohn concernem à *entrada, conteúdo, mecanismos de navegação, seleção de produtos, formulação e revisão de pedidos (ordering), carrinho de compras (shopping cart), aspectos internacionais, download, retorno de informações e recuperação de erros, ajuda online e aspectos para a atração do consumidor.*

Os avanços tecnológicos têm possibilitado aos sistemas computacionais, em especial ao hardware dedicado à renderização⁵⁶, a produção de aplicações visualmente ricas em detalhamento e perceptivelmente realísticas destinadas a ambientes virtuais (*virtual environments - VE*). No entanto, o esforço dispendido no desenvolvimento dos componentes de processos interativos em sistemas de realidade virtual tem sido relativamente muito inferior. Além disto, as interfaces raramente são avaliadas com usuários. Por conseguinte, os projetos correntes de interfaces para aplicações de realidade virtual são freqüentemente de baixa qualidade. Hix *et al.* [Hix99] argumentaram que embora a engenharia da usabilidade ainda seja uma faceta emergente em nível de desenvolvimento de sistemas desta natureza, as praticas de projeto centrado no usuário e de avaliação da usabilidade ainda deixam muito a desejar no contexto dos sistemas de realidade virtual.

Com o propósito de contribuir para a reversão deste cenário, Hix *et al.* [Hix99] e Gabbard *et al.* [Gabb99] apresentaram abordagens iterativas estruturadas de projeto e avaliação centrados no usuário para a interação usuário-sistema em realidade virtual. A aplicação sobre a qual fundamentou-se a iniciativa dos autores foi *Dragon*, um sistema de realidade virtual desenvolvido pelo laboratório de realidade virtual do *Naval Research Laboratory (NRL)* para a visualização e gestão de informações relativas a campos de batalha.

A abordagem de Hix *et al.* [Hix99] consistiu no uso iterativo de diretrizes de projeto e avaliação heurística da interface por especialistas, seguida de avaliação formativa da usabilidade e, por fim, de avaliação somativa (*summative*) do processo interativo. A abordagem de Gabbard *et al.* [Gabb99], similar em concepção à de Hix *et al.* [Hix99] e apresentada posteriormente, reflete o amadurecimento do processo na estruturação de projeto e avaliação proposta pelos autores para sistemas similares ao *Dragon*. Vale a pena mencionar que estes trabalhos se encontram entre os primeiros a relatarem a aplicação de tal abordagem em um projeto de interface para ambientes de realidade virtual.

⁵⁶ O termo *renderização*, assim como o verbo *renderizar*, vêm sendo usados na literatura brasileira sobre computação gráfica como equivalentes dos termos em inglês *rendering* e *to render*, respectivamente. *Renderizar* é gerar uma figura ou imagem bidimensional a partir de informações de cor ou nível de cinza, sombra, profundidade, etc., provenientes de um conjunto de pontos de uma representação tridimensional.

Gomes [Gome99] desenvolveu um estudo comparativo e contrastante do retorno de informações de usabilidade de estudantes de engenharia (usuários de teste) e de projetistas de aplicações multimídia, a partir da avaliação de um protótipo de *software* destinado à instrução de usuários sobre telecomunicações via recursos computacionais. Gomes combinou aplicação de questionário (tendo usado o *QUIS v. 5.5b* da *University of Maryland*) com a verbalização de procedimentos (*think-aloud*) e avaliação heurística para coletar informações sobre a usabilidade de aplicações multimídia de um universo amostral de 15 estudantes de engenharia (particionado em 3 sub-grupos: 5 americanos, 5 chineses/coreanos e 5 indianos/paquistaneses) e 5 projetistas de aplicações de multimídia educacional. O estudo realizado pelo autor fornece informações de importância para o contexto da definição de estratégias e métodos econômicos para a avaliação formativa de protótipos de educação remota (*courseware*), além de especificar instrumentos e ferramentas destinadas à avaliação de protótipos.

Virvou e Tsiriga [Virv00] evidenciaram a importância do papel representado por professores e estudantes no ciclo de vida de um sistema de tutoria inteligente, o *EasyMath*, desenvolvido para o ensino de Álgebra. Segundo as autoras, como uma das metas primárias do projeto era a utilidade e usabilidade do *EasyMath* em salas de aula, professores e estudantes de matemática foram envolvidos ao longo de todo o ciclo de vida do produto. Participaram do processo de avaliação por observação professores representando, conforme solicitação prévia dos avaliadores, o papel de um aluno médio ao interagirem com o *EasyMath*, além de 240 estudantes. Na fase posterior ao ensaio com observação, os participantes foram sondados através de um questionário elaborado pelas autoras com base nas heurísticas propostas por Squires e Preece [Squi99] para a aprendizagem fundamentada em produtos de *software*, por sua vez adaptadas daquelas propostas por Nielsen [Niel93b, Niel94a, Niel97c] para refletirem critérios sócio-construtivistas destinados à aprendizagem.

3.2.2.4 Inspeção Fundamentada na Perspectiva (*Perspective-based Inspection*)

A *inspeção fundamentada na perspectiva (perspective-based inspection)* é um método relativamente recente desenvolvido por Zhang, Basili e Shneiderman [Zhan98b, Zhan98c] (*University of Maryland*), que emergiu dos propósitos de (i) desenvolver uma nova estratégia de inspeção da usabilidade cada vez mais direcionada para uma perspectiva específica e (ii) compreender a viabilidade, efetividade e escopo de uma inspeção de usabilidade fundamentada em uma perspectiva específica.

Segundo os autores, o método desenvolvido teve com origem alguns dos resultados da pesquisa de inspeções de *software* e caracteriza-se pela integração de quatro idéias, a saber: (i) *inspeção de produtos a partir de três perspectivas diferentes*; (ii) *uso de roteiros de tarefas*; (iii) *aquisição de resultados previsíveis a partir de um processo de inspeção bem definido*; e (iv) *definição e refinamento de critérios de inspeção*.

Cada uma das três sessões de inspeção focaliza uma das perspectivas pré-definidas, a saber: (i) *uso do produto por principiantes*; (ii) *uso do produto por experientes*; (iii) *manejo de erros*. A fim de garantir uma inspeção do ponto de vista do usuário, o processo de inspeção é suportado por um conjunto de roteiros de tarefas contextualizadas e bem definidas. Cada tarefa é analisada e

decomposta em um conjunto de etapas canônicas, cada uma das quais, por sua vez, é associada a um conjunto pré-definido de critérios de inspeção. Esses critérios são definidos de modo a se adequarem ao domínio da interface inspecionada (e.g., interfaces para a Web), podendo ser refinados conforme a experiência da equipe avaliadora ou da organização que adota o método.

3.3 Considerações Finais

À medida que novas tecnologias de *hardware* e *software* emergem, suscitando novas formas de interação entre os artefatos tecnológicos e seus usuários, novas técnicas de projeto e avaliação são propostas, experimentadas e investigadas. Neste cenário, as tecnologias de compreensão da linguagem vêm sendo cada vez mais integradas às tecnologias visuais e gráficas, com o propósito de dar suporte às mais diversificadas formas de interação usuário-computador que surgem quase que diariamente.

Assim, outro vasto campo de investigação se abre para a comunidade pesquisadora de estratégias avaliatórias empíricas e métricas associadas, sobretudo porque não há denominador comum no que concerne a metodologias de avaliação e tecnologias de suporte para interfaces interativas cada vez mais complexas, nem se afigura bem definido, por conseguinte, como avaliá-las [Hirs95]. Resta apenas o senso comum da possibilidade de adaptação de alguns dos métodos de avaliação empírica existentes, mesmo que ainda não se consiga delinear precisamente a extensão e as limitações de sua aplicação a contextos a cada dia mais diversificados. Resta também a certeza (cada vez maior) da necessidade de integração de competências multidisciplinares para a proposição e investigação de metodologias avaliatórias alternativas.

É inegável a importância das metas da avaliação da usabilidade de interfaces usuário-computador. Todavia, é importante que a usabilidade seja encarada como uma síntese de diversos princípios e aspectos avaliatórios [Gui91, Swee93, Treu94, Treu98, Cons99] no domínio do desenvolvimento de *software*. Conforme tem sido tratada desde o início deste documento, a usabilidade não deve ser dogmaticamente encarada como o único fator considerado em processos de avaliação de interfaces usuário-computador, mas apenas um dos fatores mais expressivos considerados na modelagem da qualidade de produtos de *software*.

Via de regra, os métodos de avaliação da usabilidade se enquadram na categoria de métodos *diagnósticos*, uma vez que sua orientação mais corrente volta-se para a identificação, classificação e racionalização de *problemas* relativos à interface. Neste contexto, é inevitável a comparação de estratégias, de benefícios e inconvenientes inerentes a cada uma delas, assim como os julgamentos advindos de sua aplicação refletem uma comparação da interface avaliada com outras, quer em termos da quantidade de problemas identificados, quer em nível da seriedade dos problemas apresentados, como será comentado no Capítulo 3.

No entanto, há outras categorias de avaliação que também têm merecido a atenção dos profissionais da área de interação homem-máquina e, mais especificamente, da interação usuário-computador. Diversos estudos têm descrito práticas avaliatórias de caráter *descritivo*, *explanatório* ou *prescritivo*.

O emprego de *técnicas de representação comportamental e modelos analíticos* tem se mostrado uma abordagem eficiente para a avaliação de projetos de interfaces já implementadas, embora sua extensão no prognóstico do desempenho de interfaces em desenvolvimento ainda mereça ser investigado e demonstrado, conforme ressaltam Gugerty [Guge93], Treu [Treu94] e Dix et al. [Dix98]. A maioria dessas técnicas vem sendo concebida originalmente para a aplicação em projetos de sistemas interativos [Bear96, Baum00], fato que não descarta ou impede a utilização destes instrumentos em atividades de inspeção de projetos, protótipos ou produtos finais, nem tampouco atenua sua relevância enquanto fundamentos alternativos para procedimentos de avaliação de interfaces usuário-computador.

Outra modalidade bastante adotada dentro do contexto de metodologia de projeto e avaliação de interfaces usuário-computador tem sido a de estudos experimentais de *comparação de técnicas e estilos de interação* com relação a variáveis selecionadas de desempenho, empregadas na representação de amostras de usuários, e.g. Jeffries e Rosenberg [Jeff87], Svendsen [Sven91], Morgan et al. [Morg91], Eberts e Bittianda [Eber93], Zhang et al. [Zhan98d], Suhm et al. [Suhm99], Lindeman et al. [Lind99].

Experimentos avaliatórios com o propósito de comparar *características alternativas de uma técnica de interação específica*, em destaque as interações baseadas na seleção de *itens via menu* e a *manipulação direta*, também tem sido o foco de investigação de diversos pesquisadores, e.g. Mitchell e Shneiderman [Mitt89], Walker et al. [Walk91], Ballas et al. [Ball92], Gray [Gray93], Franzke [Fran95], Terwilliger e Polson [Terw96], Hornof e Kieras [Horn97], Hornof e Kieras [Horn99].

O efeito do uso de *ícones e da interação icônica no desempenho do usuário* também tem sido outro tópico a receber uma atenção crescente da comunidade pesquisadora, como atestam os trabalhos de Blankenberger e Hahn [Blan91], MacGregor [MacG92], Kaptelinin [Kapt93], Moyes [Moye94], Williams e Buehler [Will97], Brignull [Brig99].

Investigações de *efeitos de características específicas de interfaces sobre o desempenho de usuários*, envolvendo também outras técnicas gráficas, têm merecido também a atenção de vários pesquisadores, dentre os quais podem ser citados Card, English e Burr [Card78]; Epps [Epps86]; Kieras [Kier92]; Card, Pirolli e Mackinlay [Card94], Lim et al. [Lim96], Watson et al. [Wats97], Amento et al. [Amen99].

Por outro lado, diversos outros tipos de *dispositivos de entrada ou modos de entrada da informação* vêm sendo investigados: (i) *voz e toque* (Bierman et al. [Bier92], Franzke et al. [Fran93], Marx e Schmandt [Marx94], Resnick e Virzi [Resn95]; Savidis et al. [Savi96], Fuhrman e Groller [Fuhr98], Hincley e Sinclair [Hinc99a]); (ii) *dispositivos ópticos e táteis* (Goldberg e Richardson [Gold93], Kurze [Kurz96, Kurz98], Moran et al. [Mora97], Sugiura e Koseki [Sugi98], Goldstein et al. [Gold99]); e (iii) *dispositivos de entrada multidimensionais* (Jacob e Sibert [Jaco92], Rekimoto [Reki96], Balakrishnan et al. [Bala97, Bala99], Harrison et al. [Harr98], Hincley et al. [Hinc99b], Pentland [Pent00], Frölich e Plate [Fröl00]).

Enfim, inúmeros outros estudos têm se voltado mais para as relações gerais entre usuários e sistemas computacionais, sobretudo no que diz respeito à qualificação e os efeitos de tais interações (vide Palmiter e Elkerton [Palm92], Trumbly et al. [Trum93], Brewster et al. [Brew94],

Thomas e Demczuk [Thom00], Sibert e Jacob [Sibe00], Tanriverdi e Jacob [Tanr00]).

É conveniente ressaltar que todas as modalidades de avaliação citadas nos parágrafos anteriores desta seção foram revisadas no levantamento bibliográfico realizado nesta pesquisa. Não obstante, foram apenas superficialmente mencionadas e com um caráter puramente informativo neste capítulo, em virtude de não estarem diretamente relacionadas com o escopo deste trabalho.

Como último comentário, também é conveniente observar que todas as abordagens avaliatórias e técnicas associadas, descritas nas seções precedentes, representam, sem dúvida, estratégias elegíveis conforme (i) os aspectos a serem focalizados em nível da avaliação; (ii) os contextos de representação considerados para o sistema computacional analisado (e, por extensão, os modelos de interação adotados); (iii) os atributos e características nos quais se fundamenta a modelagem do usuário; (iv) e/ou até mesmo os padrões de causa e efeito que se deseja evidenciar. Não obstante, excetuando algumas das estratégias de caráter relativamente informal, e.g., as avaliações envolvendo *heurísticas*, tais técnicas tendem a apresentar um caráter mais formal, pois envolvem fatores que representam diferentes objetos da interação usuário-computador no processo de verificação de hipóteses sobre *se* e *como* um fator ou um conjunto de fatores (*variáveis independentes*) influenciam os indicadores quantitativos e/ou qualitativos considerados (*variáveis dependentes*).

Por fim, deve-se atentar para o fato de que todas as técnicas apresentadas na seção 3.2 deste capítulo têm sido desenvolvidas para dar suporte a contextos de avaliação da usabilidade de interfaces os mais diversificados, quer fundamentados no desempenho do usuário durante o uso da aplicação, quer na satisfação do usuário no tocante à interface ou na inspeção de conformidade dos mecanismos de interação a um conjunto de recomendações, heurísticas, regras ou normas. São os fundamentos do contexto avaliatório de interfaces usuário-computador, assim como as metas almejadas, que irão definir o encaminhamento da avaliação, direcionando seu foco, conforme será constatado no próximo capítulo.

4

ENFOQUES AVALIATÓRIOS DA USABILIDADE DE INTERFACES USUÁRIO-COMPUTADOR

"... I had a fervent faith that philosophy and reason would make a bridge for me by which I could go and come in both worlds."

[Anne Rice - Pandora (1999)]

4	Enfoques Avaliatórios da Usabilidade de Interfaces Usuário-Computador.....	119
4.2	Enfoques Avaliatórios Consagrados.....	120
4.3	Mensuração da Usabilidade.....	124
4.3.1	Enfoques Avaliatórios e Pontos de Vista de Processos Interativos.....	124
4.3.2	Seleção do Enfoque Avaliatório.....	127
4.4	Caráter Complementar dos Enfoques Quantitativo e Qualitativo na Pesquisa.....	128

O presente capítulo retoma a discussão de aspectos relativos à avaliação da usabilidade não explorados nos capítulos anteriores, ressaltando questões pertinentes ao contexto dos enfoques avaliatórios consagrados na literatura da área, especialmente aquelas concernentes à mensuração da usabilidade e à confrontação de enfoques quantitativos e qualitativos.

A seção 4.1 (**Introdução**) apresenta, sob a forma de representações tabulares, um estudo comparativo das diferentes estratégias avaliatórias estudadas no Capítulo 2, finalizando com uma revisão geral das categorias mais abrangentes de enfoques avaliatórios registrados na literatura da área.

A seção 4.2, **Mensuração da Usabilidade**, retoma e expande a discussão de aspectos referentes à seleção de um enfoque avaliatório, segundo diferentes pontos de vista teóricos segundo os quais são encarados os processos interativos usuário-computador, revisando características dos diferentes enfoques revisados. Adicionalmente, apresenta-se tópicos relativos aos fatores de projeto que influenciam a usabilidade de produtos de *software*, dos pontos de vista do usuário, da aplicação e do especialista.

Encerrando o capítulo, a seção 4.4 (**Caráter Complementar dos Enfoques Quantitativo e Qualitativo na Pesquisa**) estabelece uma confrontação dos enfoques revisados, justificando a adoção de um enfoque híbrido de avaliação a partir de vantagens e desvantagens de abordagens de pesquisa quantitativas e qualitativas.

4.1 Enfoques Avaliatórios Consagrados

A linha de apresentação adotada nos capítulos anteriores buscou evidenciar o caráter complementar de cada técnica com relação às demais apresentadas. Tal suplementação deve-se sobretudo ao fato de cada técnica de avaliação enfatizar aspectos distintos da engenharia da usabilidade. Além do mais, cada uma das estratégias avaliatórias anteriormente discutidas apresenta vantagens e desvantagens que as demais podem, respectivamente, incrementar e compensar. É importante observar que cada uma das técnicas apresentadas depende apenas parcialmente do número de usuários de teste disponíveis para a realização das tarefas inerentes ao contexto avaliatório ou, no caso das inspeções (de diretrizes de projeto, de padrões, de heurísticas), do número de especialistas disponíveis. O grau de experiência da equipe de avaliação também pode ser uma variável de grande influência na seleção do método ou da combinação de métodos a adotar.

Esta seção apresenta os resultados de um estudo comparativo preliminar das vantagens e desvantagens das diferentes estratégias avaliatórias estudadas, que pode servir de ponto de partida para a escolha de uma estratégia avaliatória ou de uma combinação delas que possa satisfazer às necessidades de um contexto avaliatório particular.

A seguir, são apresentados dois quadros construídos, inicialmente, a partir de informações adaptadas de compilações feitas por Jeffries *et al.* [Jeff91], Nielsen [Niel93c] e Van Vianen *et al.* [VanV96] e da revisão realizada por Queiroz e Turnell [Quei96] sobre técnicas de avaliação. Posteriormente, ambos os quadros foram revisados à luz das informações fornecidas por Dix *et al.* [Dix98], Jordan [Jord98] e Constantine e Lockwood [Cons99], as quais complementaram aquelas anteriormente contidas nos referidos quadros.

O Quadro 13 delinea um perfil sumariado das técnicas de avaliação discutidas no capítulo anterior. Embora construído de um modo bastante sucinto, a fim de oferecer um panorama sinóptico das técnicas consagradas de avaliação e facilitar a consulta, este quadro provê um acesso rápido e preciso a aspectos relevantes inerentes a cada uma das técnicas anteriormente descritas, e.g. *estágio do ciclo de vida do produto* ao qual a técnica se aplica, *número típico de usuários de teste*, *vantagens* e *desvantagens* de sua aplicação.

Adicionalmente, o Quadro 14 oferece subsídios básicos para a seleção de técnicas de avaliação, em função de sete critérios de decisão, a saber: *Recursos Necessários*, *Tempo Necessário*, *Número Mínimo de Avaliadores*, *Tipo de Registro*, *Custo Estimado*, *Classe de Problemas Encontrados* e *Percentual de Problemas Encontrados*.

Apesar de evidenciarem a necessidade de uma revisão e complementação das informações contidas em ambas as tabelas, Queiroz e Turnell [Quei96] demonstraram a aplicabilidade dos Quadros 13 e 14 através de um estudo de caso envolvendo a avaliação da interface do módulo de entrada de dados de um sistema de informações geográficas (SIG) [Quei98a, Quei98b].

Vale a pena enfatizar que se trata de uma iniciativa preliminar de concepção de um instrumento de auxílio à tomada de decisões relativas à seleção de uma técnica ou de uma associação de técnicas dentre um elenco de técnicas consagradas, cujos autores evidenciam a necessidade de revisão e complementação com graus de detalhamento e precisão maiores.

Quadro 13 – Perfil Sintetizado das Técnicas de Avaliação da Usabilidade Revisadas

Método	Estação do Ciclo de Vida	Nº de Usuários Necessários	Principais Vantagens	Principais Desvantagens	
Ensaio de Usabilidade (Usability Testing)	Observações (Observation)	<ul style="list-style-type: none"> ⊗ Análise de Tarefas ⊗ Atualização/Reprojeto 	03 ou mais	<ul style="list-style-type: none"> ⊗ Verificação da execução de tarefas reais ⊗ Facilidade de repetição ⊗ Custo relativamente baixo 	<ul style="list-style-type: none"> ⊗ Dispendio de tempo ⊗ Falta de controle da situação pelo avaliador ⊗ Influência da natureza intrusiva do ensaio no desempenho do usuário
	Questionários (Questionnaires)	<ul style="list-style-type: none"> ⊗ Análise de Tarefas ⊗ Atualização/Reprojeto 	Pelo menos 30	<ul style="list-style-type: none"> ⊗ Verificação preferências subjetivas do usuário ⊗ Facilidade de repetição ⊗ Custo relativamente baixo ⊗ Elevado potencial do público alvo 	<ul style="list-style-type: none"> ⊗ Duração total do processo ⊗ Baixa taxa de retorno ⊗ Dificuldade de elaboração ⊗ Dificuldade de análise dos resultados
	Entrevistas (Interviews)	<ul style="list-style-type: none"> ⊗ Análise de Tarefas 	Pelo menos 05	<ul style="list-style-type: none"> ⊗ Verificação aprofundada de atitudes e experiência do usuário ⊗ Flexibilidade ⊗ Detecção de problemas particulares ao usuário 	<ul style="list-style-type: none"> ⊗ Dispendio de tempo ⊗ Dificuldade de análise e comparação dos resultados ⊗ Possibilidade de condução das respostas pelo entrevistador ⊗ Contingência dos resultados à perícia do entrevistador
	Verbalização de Ações (Thinking Aloud)	<ul style="list-style-type: none"> ⊗ Projeto Iterativo ⊗ Avaliação Formativa 	03 a 05	<ul style="list-style-type: none"> ⊗ Baixo custo ⊗ Direcionamento para concepções errôneas do sistema ou da aplicação pelo usuário ⊗ Aceleração do aprendizado dos processos interativos pelo usuário 	<ul style="list-style-type: none"> ⊗ Fornecimento de resultados pouco representativos para a maioria dos processos de mensuração de desempenho ⊗ Caráter intrusivo e não "natural"
	Interação Construtiva (Constructive Interaction)	<ul style="list-style-type: none"> ⊗ Projeto Iterativo ⊗ Avaliação Formativa 	06 a 10	<ul style="list-style-type: none"> ⊗ Dinâmica do aprendizado em grupo ⊗ Verbalização mais "natural" de ações por crianças/adultos tímidos ⊗ Aquisição de um maior número de comentários do o ensaio de verbalização de ações 	<ul style="list-style-type: none"> ⊗ Dobro de usuários de teste, com relação às estratégias tradicionais de verbalização de ações ⊗ Discrepâncias comportamentais dos usuários de teste, acrescidas de possíveis diferenças de grau de aprendizagem e experiência computacional
	Discussões em Grupo (Focus Groups)	<ul style="list-style-type: none"> ⊗ Análise de Tarefas ⊗ Projeto Cooperativo 	05 a 09 por grupo Mais de um grupo	<ul style="list-style-type: none"> ⊗ Verificação de reações espontâneas ⊗ Prática de dinâmica de grupo para a detecção de problemas ⊗ Universalidade dos comentários 	<ul style="list-style-type: none"> ⊗ Dificuldade de análise das diversas variáveis envolvidas ⊗ Contingência dos resultados à perícia do moderador ⊗ Dificuldade de verbalização dos usuários experientes
	Ensaio Retrospectivo (Retrospective Testing)	<ul style="list-style-type: none"> ⊗ Análise de Tarefas ⊗ Atualização/Reprojeto 	03 ou mais	<ul style="list-style-type: none"> ⊗ Atenuação do caráter invasivo da observação direta ⊗ Dispensa de observadores ⊗ Revisão mais extensa e detalhada do comportamento dos usuários de teste ⊗ Avanço e retrocesso, além de "congelamento", de cenas previamente registradas em vídeo durante a execução dos testes de usabilidade tentados 	<ul style="list-style-type: none"> ⊗ Dificuldade de análise dos registros das sessões de teste ⊗ Pressuposição de um ensaio inicial de usabilidade ⊗ Dispendio de tempo ⊗ Inviabilidade em casos de usuários de teste que desempenham atividades críticas (impossibilidade de liberação pela empresa durante a ocorrência dos ensaios)
	Captura Automática (Automatic Logging)	<ul style="list-style-type: none"> ⊗ Testes Finais ⊗ Atualização/Reprojeto 	Pelo menos 20	<ul style="list-style-type: none"> ⊗ Continuidade do processo ⊗ Detecção de recursos da aplicação bastante usados ou muito pouco usados ⊗ Custo relativamente baixo 	<ul style="list-style-type: none"> ⊗ Violação da privacidade do usuário ⊗ Volume de dados gerados ⊗ Monitoração exclusiva de ações diretas ⊗ Contingência à confiabilidade do hardware instrumentado e do software empregado
	Retorno Imediato de Opiniões do Usuário (User Feedback)	<ul style="list-style-type: none"> ⊗ Atualização/Reprojeto 	Pelo menos 100	<ul style="list-style-type: none"> ⊗ Monitoração de mudanças nas expectativas e opiniões do usuário ⊗ Potencial de coleta de informações de interesse 	<ul style="list-style-type: none"> ⊗ Dispendio de tempo na organização e acompanhamento dos resultados ⊗ Volume de respostas e/ou informações geradas ⊗ Possível contingência a opiniões de técnicos e especialistas
Revisões Sistemáticas Cognitivas (Cognitive Walkthroughs)	<ul style="list-style-type: none"> ⊗ Início de Projeto ⊗ Ciclo Interno de Projeto Iterativo 	Pelo menos 05	<ul style="list-style-type: none"> ⊗ Auxílio à definição de metas e expectativas do usuário ⊗ Uso por engenheiros de software ⊗ Baixo custo 	<ul style="list-style-type: none"> ⊗ Necessidade de metodologia de definição de tarefas ⊗ Identificação de muito poucos problemas gerais e recorrentes ⊗ Identificação de muito poucos problemas mais sérios ⊗ Caráter tedioso e detalhista ⊗ Contingência dos resultados à perícia da equipe condutora do teste e ao entrosamento equipe avaliadora/grupo de usuários de teste 	
Padrões e Diretrizes de Projeto (Standards and Guidelines)	<ul style="list-style-type: none"> ⊗ Início de Projeto ⊗ Ciclo Interno de Projeto Iterativo ⊗ Atualização/Reprojeto 	Nenhum	<ul style="list-style-type: none"> ⊗ Identificação de problemas gerais e recorrentes ⊗ Uso por engenheiros de software ⊗ Baixo custo 	<ul style="list-style-type: none"> ⊗ Alto grau de subjetividade das recomendações ⊗ Dispendio de tempo para leitura de coletâneas volumosas de recomendações ⊗ Contingência à experiência do avaliador e à seleção do conjunto de recomendações ⊗ Identificação de muito poucas problemas mais sérios 	
avaliação Heurística (Heuristic Evaluation)	<ul style="list-style-type: none"> ⊗ Início de Projeto ⊗ Ciclo Interno de Projeto Iterativo ⊗ Atualização/Reprojeto 	Nenhum	<ul style="list-style-type: none"> ⊗ Detecção de problemas individuais de usabilidade ⊗ Identificação da maioria dos problemas ⊗ Identificação de problemas mais sérios ⊗ Baixo custo ⊗ Informalidade relativa 	<ul style="list-style-type: none"> ⊗ Alto custo ⊗ Necessidade de vários avaliadores ⊗ Não envolvimento de usuários "reais" ⊗ Escassez de avaliadores experientes em interação usuário-computador ⊗ Identificação de problemas menos sérios, passíveis de serem descartados pelos projetistas 	

Quadro 14 – Critérios de Seleção das Abordagens Metodológicas Revisadas

Método	Recursos Necessários	Tempo Necessário	Nº Mínimo de Avaliadores	Tipo de Registro	Custo Estimado (US)	Classe de Problemas Encontrados	% de Problemas por Avaliador	
Ensaio de Usabilidade (Usability Testing)	Observações (Observation)	● Lápis e papel ⊙ Gravador de áudio ⊙ Câmara de vídeo	3 semanas	02	✓ Relatório escrito ✓ Arquivo em disquete ✓ Fita de áudio ✓ Fita de vídeo	2.000,00	Generais e recorrentes graves	20
	Questionários (Questionnaires)	● Lápis e papel	8 semanas	02	✓ Relatório escrito ✓ Arquivo em disquete	2.000,00	Generais e de consistência superficiais	20
	Entrevistas (Interviews)	● Lápis e papel ⊙ Gravador de áudio ⊙ Câmara de vídeo	3 semanas	02	✓ Relatório escrito ✓ Arquivo em disquete ✓ Fita de áudio ✓ Fita de vídeo	2.000,00	Generais e recorrentes graves	20
	Verbalização de Ações (Thinking Aloud)	● Lápis e papel ⊙ Gravador de áudio ⊙ Câmara de vídeo	2 semanas	02	✓ Relatório escrito ✓ Arquivo em disquete ✓ Fita de áudio ✓ Fita de vídeo	2.000,00	Generais e recorrentes graves, intermediários e superficiais	20
	Interação Construtiva (Constructive Interaction)	● Lápis e papel ⊙ Gravador de áudio ⊙ Câmara de vídeo	2 semanas	02	✓ Relatório escrito ✓ Arquivo em disquete ✓ Fita de áudio ✓ Fita de vídeo	3.000,00	Generais e recorrentes intermediários e superficiais	20
	Discussões em Grupo (Focus Groups)	● Lápis e papel ⊙ Gravador de áudio ⊙ Câmara de vídeo	3 semanas	03	✓ Relatório escrito ✓ Arquivo em disquete ✓ Fita de áudio ✓ Fita de vídeo	2.000,00	Generais e recorrentes graves, intermediários e graves	20
	Ensaio Retrospectivo (Retrospective Testing)	● Lápis e papel ⊙ Câmara de vídeo	1 semana	02	✓ Relatório escrito ✓ Arquivo em disquete ✓ Fita de áudio ✓ Fita de vídeo	1.000,00	Generais, recorrentes e de consistência graves, intermediários e superficiais	20
	Captura Automática Direta da Aplicação (Automatic Logging from the Application)	● Software de instrumentação do sistema	4 semanas	02	✓ Relatório escrito ✓ Arquivo em disquete	1.000,00	Generais e recorrentes graves e intermediários	20
	Retorno Imediato de Opiniões do Usuário (User Feedback)	● Facilidades de correio eletrônico ● Software de coleta automática de opiniões	8 semanas	02	✓ Relatório escrito ✓ Arquivo em disquete	500,00	Generais, recorrentes e de consistência intermediários e superficiais	20
Revisões Sistemáticas (Walkthroughs)	● Lápis e papel ⊙ Gravador de áudio ⊙ Câmara de vídeo	2 semanas	03	✓ Relatório escrito ✓ Arquivo em disquete ✓ Fita de áudio ✓ Fita de vídeo	3.000,00	Generais e recorrentes intermediários e superficiais	20	
Inspeção de Conformidade (Conformance Inspection)	● Coletânea de diretrizes ● Manual de regras ● Padrão	Toda a fase de desenvolvimento ¹ 4 semanas ²	02	✓ Relatório escrito ✓ Arquivo em disquete ✓ Fita de áudio ✓ Fita de vídeo	1.000,00	Generais, recorrentes e de consistência graves, intermediários e superficiais	20	
Avaliação Heurística (Heuristic Evaluation)	● Heurística ● Lista de inspeção (Checklist)	2 semanas	02	✓ Relatório escrito ✓ Arquivo em disquete ✓ Fita de áudio ✓ Fita de vídeo	3.000,00	Generais, recorrentes e de consistência graves, intermediários e superficiais	40	

¹ Uso tradicional, como guia para desenvolvimento de projetos de interfaces.

² Uso para avaliação nos moldes de uma avaliação empírica.

Também vale a pena acrescentar que os resultados desta pesquisa servirão de respaldo para um refinamento dos aspectos sumariados no Quadro 14, uma vez que este trabalho trata da comparação de três enfoques avaliatórios e que o processo de seleção das técnicas de avaliação a serem adotadas em cada enfoque baseou-se nas informações do referido quadro.

No tocante ao enfoque avaliatório adotado, a ponderação de todos os aspectos da avaliação de interfaces discutidos nos capítulos precedentes (especialmente no capítulo anterior) e sumariados nos Quadros 13 e 14 conduz usualmente à constatação que, em geral, as iniciativas avaliatórias recaem em uma das duas grandes categorias seguintes:

(i) *avaliação centrada na interação usuário-produto*, que emprega usuários de teste reais interagindo com o produto avaliado, em ambientes reais ou laboratoriais, enquanto um avaliador ou uma equipe de avaliação coleta manual ou automaticamente informações qualitativas e/ou quantitativas inerentes ao processo interativo usuário-produto; e

(ii) *avaliação centrada na inspeção do produto por especialistas*, que se fundamenta em julgamentos do produto do ponto de vista de um avaliador ou de uma equipe de especialistas (e.g., engenheiros de usabilidade, engenheiros de *software* e de *hardware*), com base em padrões⁵⁷, diretrizes de projeto, guias de estilo, características almeçadas ou heurísticas de usabilidade, podendo ou não contar com a presença de usuários do produto.

Há ainda uma terceira categoria possível, a *avaliação híbrida* ou *adaptativa*, caracterizada pela integração das categorias mencionadas no parágrafo anterior. Analogamente ao fato de que a associação de diferentes técnicas complementares de avaliação possibilita a compensação das limitações apresentadas por cada uma delas, a hibridação dos enfoques de condução de processos avaliatórios *centrado na interação usuário-produto* e *na inspeção do produto por especialistas* também possibilita à equipe de avaliação uma análise mais abrangente do produto avaliado.

A literatura da área (e.g., Treu [Treu94], Dix *et al.* [Dix98]) também aponta uma quarta categoria de enfoque - a *avaliação centrada em modelos*, que podem ser cognitivos (e.g., GOMS [Card83], CCT [Kier85]), lingüísticos (e.g., BNF [Reis81], TAG [Payn86]), físicos (e.g., KLM [Card90], tri-estado [Buxt90]) ou de interação (e.g., PIE [Dix98]). Esta quarta categoria de enfoque avaliatório se respalda em formalismos teóricos voltados para (i) a representação de metas e tarefas do usuário, do diálogo usuário-sistema e de aspectos físicos de interação do usuário com o *hardware* (modelos cognitivos); (ii) a especificação, simulação e implementação de sistemas computacionais (modelos físicos); ou (iii) a implementação de estratégias de entrada de informações pelo usuário e de saída de resultados para o usuário (modelos de interação).

O Anexo A apresenta um diagrama que sintetiza os enfoques avaliatórios clássicos e as técnicas tipicamente adotadas em cada um deles, destacando cromaticamente tanto a categoria em que se insere esta pesquisa quanto as técnicas selecionadas para o enfoque híbrido apresentado neste documento.

⁵⁷ Conforme discutido anteriormente, os padrões podem ter uma repercussão em nível corporativo (e.g., padrões da IBM, Macintosh, Windows), nacional (e.g., padrões do DIN e ANSI ou normas da ABNT) ou internacional (e.g., padrões elaborados separada ou conjuntamente pelos comitês da ISO e IEC).

4.3 Mensuração da Usabilidade

Conforme se discutiu nos dois primeiros capítulos, os resultados advindos de processos de mensuração da usabilidade de interfaces usuário-computador são dependentes do contexto considerado, o qual pode envolver categorias significativamente diferentes de usuários, tarefas ou ambientes. Deste modo, é necessário redobrar a atenção no que diz respeito à generalização de resultados relativos a um contexto avaliatório específico, uma vez que poderão não refletir a realidade fora do contexto considerado. Além disto, os resultados advindos de estudos de curta duração da usabilidade de um produto devem ser analisados ponderada e cautelosamente, tendo em vista poderem não ter possibilitado a detecção e o registro de eventos pouco freqüentes que possam provocar impactos significativos sobre a usabilidade daquele produto, e.g. a ocorrência de erros intermitentes.

Um terceiro aspecto que não deve ser deixado de lado é o estágio de geração do produto no qual o enfoque avaliatório deverá ser adotado. Evidentemente, qualquer dos enfoques mencionados na seção anterior poderá ser adotado em qualquer dos estágios, se o termo *enfoque* for interpretado em sua conotação de mais alto nível. No entanto, adotar o enfoque *centrado no usuário* no estágio conceitual de um produto a ser desenvolvido significa empregar técnicas avaliatórias distintas daquelas que serão aplicadas no enfoque *centrado no usuário* no estágio de uso pleno do produto final [VanV96, Dix98]. Daí a importância de considerar atentamente a essência de uma descrição metodológica, ponderando se o estágio em que foi aplicada é análogo àquele em que se pretende adotá-la.

Assim, os três aspectos acima mencionados, em geral entrelaçados de um modo intrincado, reforçam a necessidade não apenas de um conhecimento das técnicas de avaliação e do estágio de desenvolvimento do produto em que se aplicam. Também é imprescindível o conhecimento do modo como estas técnicas podem ser combinadas no sentido de focalizarem (i) os princípios teóricos e os modelos que fundamentam o projeto ou o produto a ser avaliado; (ii) o especialista que inspeciona o projeto ou o produto; ou (iii) o usuário interagindo com o produto.

4.3.1 Enfoques Avaliatórios e Pontos de Vista de Processos Interativos

Sweeney *et al.* [Swee93] identificam três grandes categorias de enfoques avaliatórios de produtos interativos, a saber: (i) *centrada no usuário*; (ii) *fundamentada por princípios teóricos*; e (iii) *centrada no especialista*. Tais enfoques refletem a fonte dos dados que embasarão a avaliação. A adoção de um enfoque (ou de uma integração de enfoques) também dependerá de avanços tecnológicos e do ponto de vista adotado.

A avaliação *centrada no usuário* envolve um ou mais usuários executando tarefas em um dado ambiente. A avaliação *fundamentada por princípios teóricos* envolve a determinação de um compromisso entre um modelo do usuário e uma especificação do sistema pelo projetista, avaliador ou pesquisador e, como o termo deixa entrever, baseia-se integralmente em teorias, não envolvendo, portanto, nenhum processo interativo usuário-computador ou protótipo. A avaliação *centrada no especialista* envolve a presença de pesquisadores ou avaliadores, inspecionando o sistema de um modo mais ou menos estruturado, na tentativa de adequá-lo a critérios de projeto pré-definidos [Ross97].

Conforme mencionado na seção anterior, os diferentes processos de mensuração da usabilidade de produtos costumam ser enquadrados em uma das quatro grandes categorias de abordagens: (i) *centrada na interação usuário-produto*; (ii) *centrada na inspeção do produto por especialistas*; (iii) *fundamentada por modelos*; e (iv) *adaptativa (ou híbrida)*.

Observa-se equivalências entre o ponto de vista de Sweeney *et al.* [Swee93] e o ângulo de visão segundo o qual esta pesquisa foi desenvolvida. A abordagem *centrada no usuário* de Sweeney *et al.* equivale àquela *centrada na interação usuário-produto* descrita neste documento. A abordagem *fundamentada por princípios teóricos* corresponde àquela *fundamentada por modelos*, comentada na seção precedente. Por fim, as abordagens *centrada no especialista* (descrita por Sweeney *et al.*) e *centrada na inspeção do produto por especialistas* (descrita neste documento) são equivalentes.

Apenas a abordagem *adaptativa (ou híbrida)*, discutida anteriormente, não possui equivalente na categorização estabelecida por Sweeney *et al.* [Swee93]. E é precisamente esta a abordagem adotada nesta pesquisa, em conformidade com a definição de *usabilidade* considerada no capítulo introdutório. O propósito dos próximos capítulos é a consolidação do ponto de vista delineado neste trabalho no que diz respeito à abordagem avaliativa adotada.

No que diz respeito aos pontos de vista teóricos dos processos interativos usuário-computador, Sweeney *et al.* [Swee93] as discriminam como (i) *cognitivo*, (ii) *sócio-psicológico (ou psico-social)*, (iii) *organizacional*, (iv) *psicofisiológico* e (v) *das comunicações*.

O ponto de vista *cognitivo* atribui a base do processo interativo ao intercâmbio e processamento da informação em um contexto de raciocínio, solução de problemas e formulação de opiniões. Este ângulo de visão tem se difundido desde os anos oitenta, quando o ponto de vista dos modelos cognitivos do usuário passou a ser considerado cada vez mais importante, aliado ao argumento recíproco de que o usuário carece de um modelo apropriado do sistema [Eber97]. São exemplos as diferentes técnicas de representação do comportamento e os modelos analíticos propostos nas duas últimas décadas, conforme visto no capítulo anterior.

O ponto de vista *sócio-psicológico* focaliza o trabalho cooperativo desenvolvido por grupos de indivíduos, no contexto de processos interativos usuário-computador, através de uma interface. Este ponto de vista considera as semelhanças dos diálogos entre seres humanos e entre estes e sistemas computacionais, estando também relacionado à tendência dos seres humanos atribuírem aos sistemas que desenvolvem características humanas, como foi demonstrado por Weizenbaum [Weiz86], através do *ELIZA*, um sistema projetado para imitar um psicoterapeuta. Outros pesquisadores, e.g. Spiliotopoulos e Shackel [Spil81], concluíram que, segundo o ponto de vista sócio-psicológico, o uso de interfaces "amigáveis" (i.e. mais usáveis) proporcionavam maior satisfação aos seus usuários.

Estão inclusas neste ponto de vista as atividades de videoconferência e teleconferência, a partir das quais imagens de vídeo disputam o espaço da tela de visualização com planilhas eletrônicas, transparências, textos, etc., sendo apresentadas simultaneamente em várias janelas. Adicionalmente, os videoconferencistas podem visualizar e fazer anotações tanto de sugestões não-verbais quanto verbais. Dentre as diversas iniciativas de expansão e difusão deste ponto de vista, incluem-se os trabalhos de Gale [Gale89], Rogers *et al.* [Roge92], Cockburn e Greenberg

[Cock93], Grudin [Grud94], Grudin e Palen [Grud95], Fernandez e Perrewè [Fern95], Feldt [Feld97], Kalimo e Smith [Kali97] e Kivimäki *et al.* [Kivi00].

O ponto de vista *organizacional* envolve critérios inerentes a atividades desenvolvidas em nível organizacional na avaliação de processos interativos usuário-computador, respaldada pelo trabalho cooperativo auxiliado por sistemas computacionais (*Computer-Supported Cooperative Work - CSCW*) [Gall91, Gree92, Trau92, Vala93] através do qual estes sistemas têm adquirido uma expressividade crescente, sobretudo nos últimos cinco anos, como ferramentas de auxílio à tomadas de decisões grupais, visando o cumprimento de metas organizacionais e à solução de problemas [Baru97, Stan97a, Lain98, Tell99].

No âmbito do ponto de vista organizacional, são consagradas as iniciativas de investigação de Malone [Malo85], Gardiner [Gard86] e Christie e Gardiner [Chri90], cujos argumentos versam sobre a necessidade de inclusão de critérios organizacionais em procedimentos de avaliação de processos interativos usuário-computador, mencionando o interesse crescente em *groupware* e conceitos afins. Esta perspectiva de visão inclui a disposição e a produtividade grupal e individual (em maior ou menor grau, conforme os efeitos da estruturação da rede). Barua *et al.* [Baru97] realizaram um estudo revisivo que oferece uma visão sinóptica deste ponto de vista de projeto e avaliação de sistemas interativos usuário-computador.

O ponto de vista *psicofisiológico* fundamenta-se na abordagem de Gale e Christie [Gale87], que integra a psicofisiologia aos processos interativos usuário-computador, i.e. busca adaptar o mundo aos indivíduos, com base nas capacidades e limitações fisiológicas e psicológicas dos seres humanos. A pesquisa conduzida segundo este ponto de vista objetiva a otimização da eficácia e eficiência com que as atividades humanas são conduzidas, bem como o aprimoramento da qualidade de vida de um modo geral, tanto através do aumento da segurança, do conforto e da satisfação quanto da redução da fadiga e do estresse.

Gale e Christie [Gale87] e Stanton *et al.* [Stan97b] demonstraram benefícios específicos deste ponto de vista na interação usuário-computador, propondo um novo tipo de iniciativa de pesquisa progressista, denominada *CAFE of EVE (Controlled, Adaptive, Flexible, Experimental Office of the Future in an Ecologically Valid Environment)*, através da qual pode-se fazer mensurações simultâneas em vários domínios e validá-las através de comparações com outros domínios. Medidas ergonômicas relativas a projetos de teclados, fabricação de terminais de vídeo de baixa radiação, assim como inúmeras iniciativas direcionadas para a saúde pessoal executadas nos últimos vinte anos servem como exemplos da aplicação deste ponto de vista.

O ponto de vista *das comunicações* relaciona-se ao contexto de multimídia (particularmente à Internet e à WWW) e outros desenvolvimentos tecnológicos da informação (e.g., o apoio computacional a atividades cooperativas - CSCW), focalizando limitações que afetam o processamento e os resultados de tarefas levadas a cabo neste âmbito [Baru97, Chig97]. Ross [Ross97] cita como exemplo desta perspectiva de visão a mensuração de larguras de faixa em comunicações.

De acordo com o que foi exposto até este ponto e conforme se verificará ao longo dos próximos capítulos, esta pesquisa tende a se enquadrar no ponto de vista *psicofisiológico*, uma vez

que visa, através de um processo avaliatório, a otimização da eficácia e eficiência com que as atividades dos usuários são auxiliadas pelo produto avaliado, assim como também considera a influência da usabilidade da interface do referido produto sobre a satisfação de seus usuários.

4.3.2 Seleção do Enfoque Avaliatório

Tendo partido da caracterização de um problema⁵⁸, levando-se em consideração os avanços tecnológicos e a perspectiva adotada, conforme discutido ao longo da seção anterior, o passo seguinte no processo de seleção de um enfoque avaliatório é a tomada de decisão sobre as medidas de usabilidade a serem realizadas para auxiliar o processo de análise ou investigação intentado.

Os pontos de vista consagrados na literatura da área sobre a questão da usabilidade de produtos divergem sensivelmente, apenas no tocante às medidas de componentes da usabilidade. Grosso modo, independentemente das componentes representativas da usabilidade de um produto e do contexto considerado, a atenção do avaliador costuma se concentrar: (i) no *desempenho* do usuário ao utilizar o produto como suporte no desenvolvimento de suas atividades; (ii) na *satisfação subjetiva* do usuário com relação ao produto que o auxilia no desenvolvimento de suas atividades; e (iii) na *conformação* do produto a *diretivas* de projeto e avaliação que, implicitamente, objetivam a otimização do desempenho e da satisfação do usuário.

No que diz respeito ao *desempenho*, diversos documentos (e.g., Nielsen [Niel93c], Dumas e Redish [Duma94], ISO 9241 [ISO98], Constantine e Lockwood [Cons99]) apontam métricas/indicadores que refletem direta ou indiretamente esta componente da usabilidade de um produto. Evidentemente, o produto avaliado, os interesses ou metas almejados, o propósito da avaliação e, de um modo mais global, o contexto de uso do produto considerado definirão o modo como tais métricas/indicadores serão explorados.

A facilidade de aprendizado (*learnability*) é apontada como um dos fatores que mais afetam o desempenho do usuário de um produto, além de ser considerada por alguns pesquisadores como o atributo mais fundamental da usabilidade [Shac90, Niel93c, Pree94, Dix98]. Além destes, a efetividade (*effectiveness*), a eficiência (*efficiency*), a flexibilidade de uso (*flexibility*), a robustez (*robustness*), a funcionalidade (*functionality*) e a facilidade de memorização (*memorability*), dentre outros, são mencionados como atributos da usabilidade que afetam o desempenho do usuário com o produto [Niel93c, Duma94, Dix98, ISO98].

Quanto à *satisfação subjetiva*, inúmeros autores⁵⁹ têm abordado, ao longo dos últimos 20 anos, diferentes facetas deste atributo da usabilidade, quer em termos de proposição de instrumentos de mensuração, de sugestões e comentários e/ou da condução satisfatória de

⁵⁸ Neste contexto, o termo *problema* poderá caracterizar corriqueiramente a verificação de especificações operacionais concebidas para um produto, ao longo do processo de sua implementação, embora também se aplique à análise de aspectos particulares do processo interativo usuário-computador que mereçam atenção especial, quer na fase de desenvolvimento, quer na fase de uso pleno do produto final.

⁵⁹ Wancous e Lawler [Wano72], Swanson [Swan74], Penniman e Dominick [Penn80], Bailey e Pearson [Bail83], Ives *et al.* [Ives83], Shneiderman [Shne87], Chin *et al.* [Chin88], Estevam [Este90], Virzi [Virz91], Tetzlaff e Schwartz [Tetz91], Silva [Silv92], Nielsen [Niel93c], Nielsen e Philips [Niel93d], Nielsen e Levy [Niel93e], Queiroz [Quei94], ISO [ISO98], Constantine e Lockwood [Cons99].

ensaios de usabilidade centrados neste atributo.

Finalmente, no tocante à *conformação* do produto a *orientações* de projeto e avaliação, diversos pesquisadores têm se dedicado à compilação de coletâneas de informações cognitivas, organizacionais, técnicas e/ou experimentais - genéricas ou específicas - destinadas ao auxílio de projetistas e avaliadores de produtos. Essas informações têm sido divulgadas tipicamente sob a forma de diretrizes de projeto⁶⁰, heurísticas de avaliação⁶¹, guias de estilo [Micc95, App197] e padrões [DIN88, ISO92, ISO98], estes últimos em nível corporativo (e.g., *Apple*, *Microsoft*), nacional (e.g., *ANSI*, *DIN*, *BSI*, *ABNT*) ou internacional (*ISO*), conforme discutido no capítulo anterior.

4.4 Caráter Complementar dos Enfoques Quantitativo e Qualitativo

Qualquer elemento ou característica de natureza causal ou que implique a propiciação ou limitação/impedimento do alcance de uma meta de projeto representa um *fator de projeto*. Tome-se como exemplos de fatores de projeto as capacidades e limitações do usuário, do produto de *software* ou do sistema computacional que possibilita a interação usuário-produto, particularmente no que diz respeito à interface de usuário. Estes fatores requerem o reconhecimento e a atenção de planejadores, projetistas, executores e avaliadores de produtos, exigindo-lhes conhecimentos e habilidades extras e não triviais no que concerne aos processos de projeto, implementação e avaliação. Além do mais, tais fatores constituem "entradas" em tomadas de decisões de projeto.

A identificação, priorização, comparação, ponderação, seleção e utilização de fatores que influenciam o desenvolvimento de um produto constitui um tipo especial de "análise de fatores" [Treu94], bastante empregado por projetistas e, sobretudo, por avaliadores, principalmente nos últimos anos. À medida que os fatores de projeto selecionados e suas influências nas características do produto são estabelecidos, na qualidade de agentes propiciadores ou limitantes, estes parecem tornar-se "transparentes" aos projetistas, tendo em vista serem automaticamente incorporados e embutidos no projeto.

No entanto, esta transparência é aparente, pois esses fatores terão implicações no estágio de avaliação, pelo menos de dois modos, a saber: (i) *limitação do desempenho*, já que refletem capacidades e limitações inerentes às componentes principais do processo interativo (usuário, aplicação, interação usuário-aplicação); e (ii) *justificativa de características*, visto que tais fatores, via de regra, incluem capacidades não apenas observáveis ou mensuráveis (e.g., tempo de execução de uma tarefa), mas também escolhas (e.g., preferências expressas pelo usuário), que poderão posteriormente serem utilizadas para explicar o porquê da inclusão de uma determinada característica, em detrimento de outras.

⁶⁰ E.g., Galitz [Gal85], Smith e Mosier [Smit86b], Shneiderman [Shne87], Silva [Silv92], Mayhew [Mayh92], Queiroz [Quei94], Farias [Fari96], Constantine e Lockwood [Cons99], Nielsen [Niel00], Powell [Powe00], Almeida [Alme00].

⁶¹ E.g., Nielsen e Molich [Niel90], Nielsen [Niel92b, Niel93c, Niel94c], Muller *et al.* [Mull95], Gehardt-Powals [Gehr96], Rohn [Rohn98], Constantine e Lockwood [Cons99], Nielsen [Niel00].

Eis porque os fatores de projeto são encarados em processos interativos usuário-computador como um auxílio potencial aos procedimentos de avaliação do desempenho relativo à interação usuário-produto. Treu [Treu94] subdivide os fatores de projetos em três categorias: (i) *dados fatuais*, inerentes ao usuário, à aplicação⁶² e/ou ao computador/interface; (ii) *opiniões*, emitidas por usuários e/ou especialistas; e (iii) *dados observáveis e/ou mensuráveis*, relativos ao desempenho (vide diagrama da Fig. 10).



Fig. 10 - Agrupamento de *fatores de projeto*, segundo Treu [Treu94].

Dados sobre um usuário ou grupos de usuários são de extrema importância no contexto do projeto e avaliação de produtos interativos, quer se refiram a usuários-alvos específicos, envolvidos em um projeto, quer digam respeito a características genéricas de comunidades de indivíduos [Treu94]. Tais dados podem ser "estratificados" em diferentes subgrupos, dentre os quais vale mencionar os fatores *pessoais* (e.g., experiência, formação acadêmica, competências), *fisiológicos/físicos* (e.g., estados da visão e da audição, aptidão física), *psicológicos/cognitivos* (e.g., capacidade de memorização, compreensão) e de *interesse e uso da aplicação* (e.g., escolhas, metas, freqüência de uso).

Desde a fase de planejamento, antes da interface ser projetada de um modo sistemático e orientado ao usuário, as aplicações fundamentadas em sistemas computacionais carecem ser "dissecadas" em funcionalidades constituintes (recursos inerentes à aplicação) e padrões de funcionalidades (conjunto de diretivas que otimizam as características das funcionalidades da aplicação, assegurando a consistência e a robustez do produto). Tal atenção favorecerá a consistência do projeto em suas diferentes facetas (e.g., sintática, semântica, exploratória), além da compatibilidade entre o grau de complexidade da aplicação e as características do universo usuário para a qual será projetada.

Os fatores de projeto que refletem a natureza e a complexidade da aplicação podem ser subdivididos, segundo Treu [Treu94], em: (i) *impactos sobre o sistema computacional*,

⁶² Neste ponto, merece um comentário o uso indistinto dos termos *aplicativo* e *aplicação*, não só neste capítulo, mas ao longo de todo o documento. Na verdade, o termo *aplicativo* é um jargão empregado na área de engenharia de *software* em lugar do termo *aplicação*, no sentido de *produto aplicável* a um dado contexto.

dicotomizados em armazenamento dos dados e suporte de processamento; (ii) *impactos sobre o usuário*, discriminados em memória (humana) requerida e suporte mental de processamento; (iii) *fatores interativos*, relativos às facilidades necessárias à entrada de dados por parte do usuário e à saída de dados computacional; e (iv) *fatores de status e de propósitos*, que dizem respeito ao *status* da aplicação (e.g., se já existente ou apenas em fase de planejamento) e aos propósitos aos quais se destina (e.g., edição de textos, processamento de informações geográficas).

Dados computacionais fatuais devem refletir *se e como* poderão servir de suporte às aplicações residentes (dados fatuais da aplicação), a serem acessadas pelo usuário (dados fatuais do usuário) através da interface, conforme estipulações de desempenho estabelecidas a partir de contextos de trabalho. Os dados fatuais relativos à interação podem ser discriminados em: (i) *fatores de software*, relativos a linguagens de programação, sistemas operacionais, etc.; (ii) *fatores físicos/de hardware*, que incluem diferentes aspectos do *hardware*, tais como velocidade da CPU, capacidade de armazenamento, arquitetura considerada, etc.; (iii) *fatores da interface*, que integram aspectos de *software* (e.g., linguagem de interação) e de *hardware* (e.g., dispositivos interativos); e (iv) *fatores relativos à conformidade com a aplicação*, relativos aos tipos de aplicações para as quais o sistema computacional se adequa bem ou não.

Por fim, são as opiniões de especialistas e os resultados de testes de desempenho, convenientemente planejados e conduzidos com base no usuário-alvo e/ou na aplicação-alvo, que estabelecem o equilíbrio, em nível do desenvolvimento (ou do reprojeto) do produto, entre os fatores inerentes ao usuário, à aplicação e/ou ao computador/interface. Obviamente, a introdução de novas tecnologias no mercado de consumo (e.g., interfaces de visualização múltipla, multimídia, interfaces de voz e gestos) e novas concepções de trabalho (e.g., CSCW) implicam a inserção de novos fatores no contexto do desenvolvimento de produtos interativos.

É um fato que os resultados da mensuração do desempenho representam papéis preponderantes tanto durante o estágio de projeto quanto no estágio de avaliação contínua de um produto, desde a prancheta até o uso pleno [Niel93c, Treu94, Duma94]. Além do que já foi descrito sobre a mensuração do desempenho em seções anteriores, diversos autores detalham diferentes aspectos do ato de mensurar (Nielsen [Niel93c], Treu [Treu94], Dumas e Redish [Duma94], Queiroz e Turnell [Quei96], Dix *et al.* [Dix98]), o que explicita a grande significância das medidas para diversas metodologias de avaliação de produtos interativos usuário-computador.

A definição e mensuração de indicadores quantitativos no contexto da interação do usuário com um sistema (e.g., frequência de uso de comandos, tempos de execução de tarefas, número de erros repetidos) podem refletir tanto a complexidade do projeto da linguagem de interação quanto o conhecimento dessa linguagem por uma determinada comunidade usuária, a partir da análise comparativa com outros usuários mais ou menos experientes. Por outro lado, os indicadores quantitativos usualmente não oferecem resultados explícitos da satisfação do usuário no tocante às características da interface e às facilidades que esta lhe oferece. No entanto, satisfazer ao usuário constitui uma das metas primárias no âmbito do projeto de interfaces.

Durante a década passada, muita atenção foi dada ao conceito de produto "amigável". Em parte, devido às dificuldades de determinar se os usuários estavam realmente satisfeitos com as interfaces implementadas. Os setores de publicidade de diversos desenvolvedores de sistemas

computacionais passaram então a se referir às interfaces desenvolvidas como "amigáveis com o usuário". Assim, o termo *amigável* tornou-se sinônimo de características positivas do *hardware* e *software* de interface, não obrigatoriamente associadas ao usuário, implicando condições favoráveis, amigáveis, encorajadoras e não hostis do produto para com o usuário.

Mesmo hoje em dia, quando o termo perdeu a popularidade de uso, ainda presume-se que se uma interface é amigável, esta satisfará as necessidades do usuário. Porém, um sistema computacional pode ser projetado para apresentar inúmeras características "amigáveis" e, mesmo assim, não satisfazer ao usuário no tocante à realização de seus propósitos. Isto se assemelha a lidar com alguém bastante amigável que, apesar de apresentar algumas qualidades recomendáveis, pode causar irritação e, até mesmo, raiva em um interlocutor. Portanto, a qualidade "amigável" um sistema não deve ser encarada como uma condição *sine qua non* para a satisfação do usuário. No entanto, poderá ser considerada como uma causa ou fator contribuinte de sua satisfação. Talvez por esta razão os pontos de vista de estudo da satisfação do usuário tenham gerado mais controvérsia do que concordância. Kim [Kim89, in ISWN98] discriminou em sua revisão bibliográfica pelo menos três linhas de estudo, a saber: (i) satisfação em termos de atitudes do usuário com relação a sistemas de informação; (ii) satisfação em termos da qualidade da informação; e (iii) satisfação em termos da eficácia de sistemas de informação.

Os esforços de aquisição de dados qualitativos associados a produtos interativos têm se concentrado na *satisfação subjetiva* (e.g., Bailey e Pearson [Bail83], Ives, Olson e Baroudi [Ives83]), no *uso* (e.g., Lucas [Luca75], Baroudi, Olson e Ives [Baro86]), no *envolvimento do usuário* (e.g., Barki e Hartwick [Bark89]) e na *aceitação do produto pelo usuário* (e.g., Davis [Davi93]). A mensuração do sucesso de produtos interativos respaldada na satisfação de usuário tem se tornado, pouco a pouco, um procedimento corriqueiro e usual (vide Ives e Olson [Ives84] e Delone e McLean [Delo92]), tendo sido concebidos e validados diversos instrumentos padronizados para este propósito⁶³. Entretanto, a maioria desses instrumentos é direcionada para a avaliação de aplicações específicas, ao invés de poderem ser aplicados a problemas de cunho mais genérico ou de serem adaptáveis a diferentes contextos de estudo.

Um terceiro pilar no âmbito das metodologias atuais destinadas ao desenvolvimento e avaliação de produtos de *software* tem sido erguido em função de abordagens "normativas" (e.g., ISO9126 [ISO92], ISO 9241 [ISO98]). Boa parte da pesquisa metodológica orientada à qualidade tem sido destinada à geração de ferramentas e técnicas de avaliação e validação da robustez e confiabilidade do produto. A partir da perspectiva descortinada por tais atividades de pesquisa, a qualidade do produto é mensurada em função de diferentes fatores, dentre os quais: (i) a identificação e redução progressiva de omissões e falhas, durante as fases de desenvolvimento do produto e, de um modo mais abrangente, de todo o seu ciclo de vida (aceita-se comumente que *sempre* há espaço para a otimização de qualidade); (ii) o planejamento, a monitoração e as modalidades de controle das fases de desenvolvimento (i.e., o "processo" de desenvolvimento do produto). No âmbito dessas abordagens "normativas", as questões relativas à usabilidade ocupam um nicho próprio, equivalendo a um capítulo nos planos de avaliação da qualidade do produto (vide ISO 9241 [ISO98] e também o MUSiC, uma das mais recentes iniciativas focalizadas na

⁶³ E.g., Bailey e Pearson [Bail83], Shneiderman [Shne87], Kirakowski [Kira87], Kirakowski e Corbett [Kira88], Baroudi e Orlikowski [Baro88], Chin *et al.* [Chin88], Doll e Torkzadeh [Doll88], MacLean [MacL91], Lewis [Lew91], Kirakowski e Corbett [Kira90], Kirakowski [Kira93, Kira94]

usabilidade de produtos de *software* e originado de aproximações normativas) [Visc97].

Embora estas abordagens tenham representado um empreendimento relevante ao desenvolvimento e avaliação de produtos interativos, não representam adequadamente a grande variedade de filosofias, metodologias, métodos e técnicas vigentes entre as comunidades tecnológicas de informação norte-americanas e européias. Basta analisar a área de concentração denominada *projeto de software (software design)* [Wino95], que vem se tornando um recipiente para as contribuições advindas das áreas de *projeto centrado no usuário* (vide e.g., Norman e Draper [Norm86]) e *projeto participativo* (vide e.g., Greenbaum e King [Gree91]), respectivamente (ver também [HCI96]). O propósito principal da área de *projeto de software* é propiciar a emergência de uma nova perspectiva em desenvolvimento e programação de *software*, que possa oferecer uma contribuição original ao levantamento e à satisfação das necessidades cognitivas de usuários finais e das condições de trabalho circunstanciais e operacionais [Visc97].

O panorama descortinado segundo estes três pontos de vista distintos, embora apresentando zonas de interseção, sugere a investigação de aspectos referentes à complementaridade dos resultados obtidos a partir de cada um deles, às possíveis discrepâncias e similaridades dos modos como cada um deles aborda os mesmos aspectos do problema considerado, às possíveis soluções de compromisso entre os diferentes pontos de vista.

O argumento corriqueiro sobre os méritos relativos às linhas de pesquisa que, de um modo geral, são rotuladas como qualitativa e quantitativa costuma ser distorcido por dois fatos: (i) a falta de definições coerentes; e (ii) o direcionamento do enfoque da maioria das discussões para os métodos, ao invés de sê-lo para as suposições básicas das posturas ontológicas e epistemológicas. Wildemuth [Wild93] sugeriu que a diferença que existe entre os paradigmas positivistas e interpretativos reside no fato daqueles reconhecerem uma realidade objetiva independente do avaliador, enquanto estes visualizam a realidade como algo construído subjetiva e socialmente. A autora associa a suposição epistemológica ao método, embora visualize a determinação do método a partir da epistemologia, não o contrário.

Neste sentido, Olson [Olso94] considerou o direcionamento do foco da maioria das discussões para o método como o cerne da distorção dos méritos atribuídos às abordagens quantitativa e qualitativa, encarando a falta de definições coerentes como uma das manifestações de tal distorção. Em seu artigo, Olson realizou um estudo da conotação dada aos termos *quantitativo* e *qualitativo* por diversos pesquisadores. Quanto às diferenças subjacentes dos paradigmas de pesquisa (objetivo *versus* subjetivo), a autora sugere que se deve focalizar principalmente a clareza dos pontos de vista ontológico e epistemológico, a qual possibilitará a seleção da abordagem metodológica mais apropriada para a solução do problema investigado.

A partir de uma revisão da literatura da área, verifica-se que o enfoque quantitativo apresenta, pelo menos, quatro grandes vantagens sobre o qualitativo: (i) é apropriado para a mensuração objetiva de diversos aspectos relativos ao uso de produtos por comunidades usuárias; (ii) é robusto no que diz respeito à mensuração de aspectos descritivos de processos interativos; (iii) propicia a realização de análises comparativas e adaptações e reproduções em outros contextos de estudo; e (iv) oferece graus de confiabilidade e eficácia possíveis de serem comprovados de forma mais objetiva.

Em contrapartida, as limitações de tal enfoque referem-se principalmente às falhas no processo de averiguação mais aprofundado de significados e explicações subjacentes, até mesmo quando significativos, fidedignos e válidos. Grosso modo, o enfoque quantitativo "reduz os participantes a um conjunto de variáveis que são de algum modo equivalentes independentemente dos indivíduos e situações consideradas" [Reas81]. A pesquisa quantitativa é poderosa na mensuração de tais variáveis. Se a mensuração sob este prisma for o foco da pesquisa, então a adoção de uma abordagem quantitativa é plenamente justificada.

Todavia, os contextos avaliatórios de produtos interativos também envolvem fatores psicológicos, como anteriormente comentado, tais como sentimentos e cognição, na maioria das vezes tão importantes para as análises quanto a mensuração de parâmetros. Embora possa ser adotado um enfoque quantitativo para mensurar tais fatores qualitativos, este apresenta usualmente grandes limitações no sentido de explicá-los mais a fundo. Uma limitação adicional do enfoque quantitativo refere-se às tendências de captar "instantâneos" de uma situação, a partir da mensuração de variáveis em instantes de tempo específicos. Muitos processos interativos usuário-sistema são afetados por variações temporais, e.g. a aquisição de experiências com relação às funcionalidades do produto ou tendências de aceitação ou rejeição do produto, nem sempre identificáveis com base em um único estudo quantitativo.

Por outro lado, o enfoque qualitativo apresenta as seguintes vantagens sobre o quantitativo: (i) permite a exploração de componentes cognitivas e afetivas em maior profundidade; (ii) encoraja o informante a introduzir conceitos importantes do ponto de vista do usuário, ao invés de aderir a aspectos pré-determinados pelo investigador, sendo mais flexível e, portanto, apropriado para a pesquisa de natureza exploratória; e (iii) possibilita a identificação de variações longitudinais das variáveis de interesse, enquanto que a abordagem quantitativa tende a fazer um "instantâneo" do comportamento, cognição ou sentimento, no momento em que a pesquisa é efetuada.

Entretanto, algumas objeções à abordagem podem ser constatadas através da leitura da documentação da área. O principal argumento contra o enfoque qualitativo refere-se à validade das descobertas e inferências, sobretudo no que concerne à dificuldade de determinação da veracidade das conclusões a que o enfoque conduz. Os números relativamente baixos associados às amostras consideradas também podem conduzir a críticas sobre a não representatividade da população [Jone97].

A objetividade de pesquisa quantitativa tem sido erroneamente interpretada como sinônimo de "pesquisa de boa qualidade", enquanto a falta de objetividade inerente à pesquisa qualitativa a tem relegado equivocadamente à categoria de "pesquisa malfeita" [Mayk94] e "não científica" [Nau95]. Adicionalmente, a adoção de um único enfoque tem sido defendida por vários autores, embora muitos dos argumentos nos quais se fundamenta sejam decididamente pragmáticos, tais como restrições temporais, necessidade de delimitação do escopo dos estudos e a dificuldade de publicação dos resultados [Cres94].

O Quadro 15 estabelece uma confrontação entre os enfoques qualitativo e quantitativo, apresentando características relevantes de cada um deles. Observe-se o caráter complementar de ambos os enfoques.

Quadro 15 - Confrontação de Enfoques Metodológicos

ENFOQUE QUANTITATIVO	ENFOQUE QUALITATIVO
↳ Hipóteses tipicamente precisas, formuladas no início da pesquisa	↳ Hipóteses tipicamente emergentes durante o desenvolvimento da pesquisa
↳ Definições tipicamente precisas, estabelecidas no início da pesquisa	↳ Definições tipicamente contextuais ou estabelecidas conforme o progresso da pesquisa
↳ Dados tipicamente reduzidos a números, com a atenção focalizada na análise e otimização da confiabilidade das contagens	↳ Descrições tipicamente narrativas, sob a forma de inferências, cuja confiabilidade é considerada adequada ao contexto
↳ Análise da validade do processo respaldada por índices estatísticos	↳ Análise da validade do processo respaldada por triangulação de dados
↳ Técnicas preferencialmente aleatórias no processo de obtenção de amostras significativas	↳ Amostras preferencialmente intencionais
↳ Preferência por controle estatístico de variáveis extrínsecas	↳ Preferência por descrições narrativas de procedimentos e análise lógica no controle/contabilidade de variáveis extrínsecas
↳ Síntese preferencialmente estatística dos resultados	↳ Descrição preferencialmente narrativa dos resultados
↳ "Dissecação" preferencial de fenômenos complexos em partes específicas para análise	↳ Descrição preferencialmente holística de fenômenos complexos

O aspecto-chave da pesquisa de metodologias híbridas reside no fato de que ambos os enfoques metodológicos, quantitativo e qualitativo, apresentam, quando considerados isoladamente, capacidades e limitações. O processo de hibridação poderá, sem dúvida, contribuir para ressaltar suas contribuições mais relevantes. O avaliador deverá buscar atingir a condição a partir da qual a combinação dos enfoques qualitativo e quantitativo possam conduzir a resultados que destaquem os pontos positivos de ambos [Nau95], onde os dados qualitativos possam respaldar e esclarecer o significado da pesquisa quantitativa [Jaya93].

As suposições seguintes deverão assegurar ao avaliador a maximização as potencialidades inerentes a uma abordagem metodológica híbrida:

- (i) as técnicas qualitativas, em especial a observação e as entrevistas abertas, possibilitam ao pesquisador o desenvolvimento de um "panorama" global do tema em investigação;
- (ii) análises quantitativas podem se afigurar mais apropriadas para a avaliação comportamental ou descritiva do contexto em que a pesquisa se insira;
- (iii) análises descritivas podem propiciar a extração de uma amostra representativa para a análise qualitativa, tendo em vista que a pesquisa quantitativa pode confirmar ou negar a representatividade de um grupo de amostras para tal estudo qualitativo;
- (iv) a maioria dos processos interativos usuário-computador envolvem aspectos cognitivos e afetivos, assim como questões comportamentais, o que justifica a adoção de um enfoque qualitativo para o estudo de aspectos dessa natureza, segundo o ponto de vista dos informantes;

(v) grande parte da pesquisa da área ainda é exploratória, fato que reforça o uso de técnicas qualitativas, com propósitos de fomentar a ocorrência de descobertas inesperadas;

(vi) análises quantitativas podem complementar as descobertas advindas da adoção de técnicas qualitativas;

(vii) análises quantitativas podem confirmar ou refutar quaisquer dados aparentemente significativos que advenham do processo avaliatório; e

(viii) a inclusão de técnicas e análises quantitativas no âmbito da pesquisa de processos interativos pode aumentar a probabilidade de divulgação dos resultados, especialmente em periódicos de tendências significativamente positivistas.

Assim, a adoção de uma metodologia fundamentada na hibridação dos enfoques quantitativo e qualitativo poderá possibilitar ao avaliador indicações sobre a representatividade do universo amostral pré-definido na fase de planejamento do enfoque qualitativo, a partir dos requisitos objetivos estabelecidos na fase de planejamento do enfoque quantitativo.

Embora a maioria dos investigadores desenvolva trabalhos de pesquisa quantitativos ou qualitativos, alguns investigadores têm sugerido associações de um ou mais enfoques metodológicos em seus contextos de trabalho, justificando processos de "triangulação" de resultados, como ocorre com Ragin [Ragi87], Kaplan e Duchon [Kap188], Lee [Lee91] e Gable [Gabl94]. Um exemplo empírico do uso de triangulação é o documento de Markus [Mark94] sobre correio eletrônico.

O enfoque avaliatório adotado na pesquisa ora descrita envolve, conforme se deixou entrever no desenrolar desta seção, resultados quantitativos e qualitativos, visto lidar tanto com mensurações, estimativas e sínteses numéricas quanto com aspectos descritivos e holísticos do alvo de estudo.

Os dois próximos capítulos tratarão das diferentes facetas do enfoque avaliatório considerado no âmbito desta pesquisa. No Capítulo 5, serão apresentados e comentados (i) aspectos referentes à importância e aos benefícios advindos do uso de padrões em processos avaliatórios; (ii) o padrão internacional *ISO 9241* e suas partes; e (iii) o procedimento metodológico adotado na inspeção de conformidade do produto considerado neste trabalho (o *MATLAB V. 5.3*) com as Partes 11, 14, 15 e 16 do padrão internacional *ISO 9241*.



INSPEÇÃO DE CONFORMIDADE DE PRODUTOS COM O PADRÃO INTERNACIONAL ISO 9241

"The metrification process has gone on for 200 years and it is still going forward... it doesn't really matter if it is this year or next year, but only that you go in the right direction."

[Anders Thor - Secretário do ISO/TC 12 e ISO/TC 203]

4	Inspeção de Conformidade de Produtos com o Padrão ISO 9241	137
5.1	Importância e Benefícios Advindos da Padronização	139
5.2	Aspectos da Padronização da Interação Homem-Máquina	145
5.3	Padrões Destinados à Interação Homem-Máquina	148
5.4	O Padrão ISO 9241	153
5.5	Metodologia Adotada na Inspeção do Produto-Alvo com Base nas Partes 11, 14, 15 e 16 do Padrão ISO 9241	158
5.5.1	Aplicação das Partes 14, 15 e 16 do Padrão ISO 9241	160
5.5.1.1	Adequação de Diálogos via Menus, Linguagens de Comandos e Manipulação Direta	160
5.5.1.2	Aplicabilidade das Recomendações da ISO	162
5.5.1.3	Síntese dos Aspectos Interativos Abordados pelas Partes 14, 15 e 16 do Padrão ISO 9241	163
5.5.2	Considerações sobre a Aplicabilidade e Adoção das Partes 14, 15 e 16 do Padrão ISO 9241	164
5.5.2.1	Avaliação de Produtos	164
5.5.2.2	Aplicabilidade	165
5.5.2.3	Métodos para a Determinação da Aplicabilidade e/ou Adoção das Recomendações	166
5.5.3	Uso das Listas de Inspeção (Checklists) das Partes 14, 15 e 16	169
5.5.4	Sumário dos Resultados de uma Inspeção	171
5.6	Comentários Finais	172

Este capítulo discute aspectos relativos ao procedimento metodológico adotado na inspeção do produto considerado nesta pesquisa – o **MATLAB v. 5.3** (The MathWorks Inc.) – com base nas Partes 11, 14, 15 e 16 do padrão internacional ISO 9241.

Na seção 5.1 (Importância e Benefícios Advindos da Padronização), discute-se a questão da padronização no contexto dos processos interativos usuário-computador, enfatizando a importância e benefícios advindos do uso de padrões em atividades de projeto e na avaliação de interfaces.

A segunda seção (**Aspectos da Padronização da Interação Homem-Máquina**) apresenta alguns prós e contras das iniciativas de padronização da interação homem-máquina, encaminhando a discussão para a apresentação dos padrões atualmente disponíveis e em desenvolvimento relacionados com diferentes aspectos dos processos interativos homem-máquina, feita na seção 5.3 (**Padrões Destinados à Interação Homem-Máquina**).

A seção 5.4 - **O Padrão ISO 9241** – apresenta com maior grau de detalhamento o padrão ISO 9241, destacando as *Partes 11, 14, 15 e 16*, utilizadas neste trabalho.

A penúltima seção (**Metodologia Adotada na Inspeção do Produto-Alvo com Base nas Partes 11, 14, 15 e 16 do Padrão ISO 9241**) detalha as etapas do procedimento metodológico adotado no processo de inspeção da conformidade do **MATLAB v. 5.3.0** (*The MathWorks Inc.*) às *Partes 11, 14, 15 e 16* do padrão internacional ISO 9241.

Encerrando o capítulo, a seção 5.6 (**Comentários Finais**) estabelece um elo de ligação com os próximos capítulos deste documento.

5.1 Importância e Benefícios Advindos da Padronização

Como se sabe, a usabilidade de um produto de *software* reflete diferentes aspectos da organização que o desenvolveu, e.g. a natureza, a estrutura e a cultura da organização, o grau de maturidade do processo de gestão das atividades de projeto e avaliação, os métodos e modelos adotados no processo de desenvolvimento. Tais aspectos, aliados aos valores e regras operacionais cultivados pela organização, podem exercer uma influência significativa sobre a efetividade da aplicação de estratégias de projeto centradas no uso, conforme visto no Capítulo 2.

Conforme a apresentação dos capítulos anteriores, também se sabe que a otimização da usabilidade de produtos de *software* pode ocorrer através de muitas estratégias, todas elas direcionadas para o desenvolvimento do produto e envolvendo atividades de avaliação de sua interface com o usuário. Neste sentido, as abordagens de diversas organizações para a usabilidade dos produtos que desenvolvem tendem a evoluir através da definição de regras e do estabelecimento de procedimentos bem estruturados. É no contexto dessas abordagens que se insere a elaboração de guias de estilo ou a adoção de padrões destinados ao desenvolvimento de produtos ergonômicos.

O termo *padrão* tem sido empregado em acepções as mais variadas e em sentidos freqüentemente confusos. Segundo a ISO [ISO99b], *padrões* são acordos documentados contendo especificações técnicas ou outros critérios precisos a serem consistentemente adotados como regras, diretrizes ou definições de características, a fim de assegurarem a adequação de materiais, produtos, processos e serviços aos propósitos a que se destinam. A IEC [IEC00] complementa a definição da ISO [ISO99b], afirmando que um *padrão* é um documento estabelecido por consenso e aprovado por um corpo reconhecido, que oferece, para uso comum e repetido, regras, diretrizes ou características de atividades ou de seus resultados, almejando atingir o estado ótimo de ordem em um dado contexto.

Neste documento, o termo assume a conotação de conjunto de especificações técnicas de atributos, processos, produtos e serviços, desenvolvida a partir de cooperação voluntária e publicamente disponível para uso em contextos pertinentes. Esta conotação é consonante com a visão considerada nos processos de padronização em nível de organizações ou corporações nacionais (e.g., ANSI, DIN, BSI, ABNT) ou internacionais (ISO, IEC), cujos documentos são normalmente descritos como consensuais. Tal conotação também engloba a visão de Smoot [Smoo95], incluindo especificações técnicas elaboradas por desenvolvedores do mundo inteiro para seus próprios produtos, processos, procedimentos e/ou serviços.

Outro esclarecimento pertinente diz respeito à visão simplista da maioria dos indivíduos sobre os objetivos fundamentais da padronização. Padronizar não significa apenas minimizar variações desnecessárias. Desta perspectiva, haveria idealmente um padrão para qualquer categoria de produto. Por conseguinte, a conformação daquele produto a um determinado padrão daria ao proprietários ou usuário alguma indicação sobre a qualidade do produto. Entretanto, há outra razão bem mais significativa para a existência da variedade de padrões correntes, em especial no que diz respeito às questões relativas ao projeto e avaliação de interfaces com os usuários. É necessário ter em mente que a tecnologia computacional embasa muitas indústrias de naturezas as mais diversas e, nesse

cenário, os padrões podem ter um impacto relevante no sucesso de mercado.

A adoção de padrões não harmônicos para tecnologias similares por diferentes países ou regiões pode contribuir para o estabelecimento de “barreiras técnicas” nos processos de exportação e importação, comprometendo a dinâmica do mercado internacional. A tomada de consciência de corporações industriais de todo o mundo no tocante à necessidade de elaboração de padrões voltados para a racionalização das transações de exportação e importação (e a conseqüente otimização do fluxo internacional dos negócios) implicou a criação de uma organização para a padronização em nível internacional, a *International Organization for Standardization* (ISO⁶⁴).

Na verdade, a padronização em nível internacional teve início em 1906, na área da Eletrotécnica, com a criação da IEC - *International Electrotechnical Commission*. Em outras áreas, porém com maior ênfase na Engenharia Mecânica, são pioneiros os esforços dispendidos para a padronização internacional pela *International Federation of the National Standardizing Associations* (ISA), criada em Nova York em 1926 e desativada em 1942, devido à II Guerra Mundial. Entre 14 e 26 de Outubro de 1946, representantes de 25 países (65 delegados) reuniram-se em Londres, em uma conferência de organizações para a padronização em nível nacional, na qual foi tomada a decisão da criação de uma nova organização para a padronização em nível internacional, cujo objeto de ação seria a facilitação da coordenação e unificação de padrões industriais.

A nova organização, nascida da união da *International Federation of the National Standardizing Associations* (ISA) com o *United Nations Standards Coordinating Committee* (UNSCC)⁶⁵, foi denominada *International Organization for Standardization* (ISO). A ISO iniciou oficialmente suas atividades em 23 de Fevereiro de 1947, tendo publicado o primeiro padrão em 1951, com o título ***Standard reference temperature for industrial length measurement***.

Nos dias atuais, a padronização no âmbito internacional já se encontra consolidada para várias tecnologias associadas a diversas áreas (e.g. processamento da informação e comunicações, indústrias têxteis, produção e utilização de energia, serviços bancários e financeiros), tendendo a crescer em importância para todos os setores de atividades industriais. Há diversos fatos que corroboram ambas as afirmações, dentre os quais a ISO [ISO99b] destaca:

(i) o *progresso do processo de globalização*, que vem estimulando cada vez mais a expansão de mercados, o que demanda o desenvolvimento de acordos documentados consensuais de reconhecimento internacional que possam servir como linguagem de negócios entre parceiros multinacionais;

(ii) a *interpenetração de setores industriais*, que acentua a demanda da normatização de componentes, produtos, regras de aplicação, etc., não somente em nível de setores similares da indústria em diferentes países ou regiões, mas também em nível dos diferentes setores interagentes no contexto multinacional;

⁶⁴ Segundo documento da ISO [ISO99b], a forma reduzida ISO não é um acrônimo de forma extensa *International Organization for Standardization*, como muitos imaginam. ISO é, de fato, um nome curto, derivado do prefixo grego *isos*, que significa *igual*, evocando uma linha de pensamento que conduz facilmente à palavra *padrão*.

⁶⁵ Criado apenas em 1944 e administrado em Londres.

(iii) a *globalização dos sistemas de comunicação*, que implica a carência urgente e progressiva de padronização em nível mundial, visando a plena compatibilidade entre sistemas abertos e, por conseguinte, estimulando a competitividade entre os produtores e multiplicando as opções para os consumidores;

(iv) a *emergência de novas tecnologias* em diversas áreas (e.g., materiais avançados, meio ambiente, ciências biológicas, urbanização e construção), demandando a definição da terminologia e a consolidação de bases de dados referentes a informações quantitativas, sobretudo em estágios de desenvolvimento em que as aplicações podem ser vislumbradas, porém não há ainda protótipos funcionais para validá-las;

(v) a *existência de países em desenvolvimento*, que suscita o apoio de infra-estruturas de padronização como uma condição primordial para o sucesso de políticas econômicas com fins a um desenvolvimento auto-sustentável, atingido mediante a otimização da produtividade, o grau de competitividade nos mercados e a capacidade de exportação.

Neste contexto, o estágio de padronização industrial em larga escala é uma condição passível de ser atingida em setores industriais específicos nos quais a grande maioria dos processos, produtos e/ou serviços são conformes aos mesmos padrões. Tal condição resulta de acordos consensuais estabelecidos entre todos os protagonistas daquele setor industrial, i.e., fornecedores, consumidores e, via de regra, governos, no que diz respeito a especificações e critérios a serem consistentemente aplicados na seleção e classificação de materiais, na manufatura de produtos e na provisão de serviços.

De um modo abrangente, os principais objetivos das iniciativas destinadas à padronização, quer em nível internacional, regional ou nacional, é a operacionalização dos processos de exportação e importação, o intercâmbio de conhecimentos e a transferência de tecnologias através da (i) *maximização da qualidade e confiabilidade dos produtos paralelamente à minimização da relação custos/benefícios*; (ii) *a melhoria da qualidade de vida, através de normas relativas às condições sanitárias, à segurança e à preservação do meio ambiente, com ênfase na minimização do desperdício*; (iii) *elevação da compatibilidade e interoperabilidade de bens e serviços*; (iv) *simplificação para a otimização da usabilidade*; (v) *redução do número de modelos e, por conseguinte, dos custos associados à sua concepção*; (vi) *maximização da eficiência na oferta de bens e facilidade de manutenção dos bens adquiridos*; (vii) *aumento da confiança dos usuários em produtos e serviços oferecidos no mercado de consumo, mediante o estabelecimento de mecanismos legais de proteção ao consumidor* [Stew98, ISO99b, Stew00].

Nos dias atuais, constata-se a elevação do nível de confiança dos consumidores em produtos implementados e/ou serviços prestados em conformidade com padrões. Nos EUA, por exemplo, as visões industriais sobre padrões e graus de conformidade são apresentados em um excerto de um relatório do *National Research Council*, intitulado *Standards, Conformity Assessment, and Trade into the 21st Century* [NRC95]. O relatório ressalta o papel dos padrões como agentes facilitadores da comercialização de produtos, informando aspectos de produtos e serviços aos consumidores previdentes, enfatizando que tal propósito não poderia ser cumprido se os consumidores não tivessem nenhum modo para determinar se as reivindicações dos fabricantes por conformidade a padrões eram corretas e justificadas.

Certificações de conformidade são termos de compromisso assumidos pelos desenvolvedores com os usuários tanto através de declarações de garantia de seus produtos quanto pelas especificações de características passíveis de inspeção por entidades independentes credenciadas junto a organizações e corpos de padronização para a condução de análises de conformidade de produtos. O termo *análise de conformidade (conformity assessment)* abrange as medidas executadas tanto pelos fabricantes, quanto por seus clientes, autoridades de normatização e terceiros independentes, com o propósito de avaliar a conformidade de produtos a padrões. São integrantes de um sistema de análise de conformidade os testes de produtos, a certificação e os serviços laboratoriais de aprovação de produtos.

Análises de conformidade e padrões estão intimamente relacionados, tendo em vista que as análises dependem da existência de padrões, ao mesmo tempo que lhes valorizam a existência. Padrões isentos de ambigüidade possibilitam o julgamento de produtos, processos e serviços para os quais foram elaborados a partir de análises de conformidade, proporcionando aos usuários a garantia de que os produtos estão em conformidade com os padrões vigentes, ao mesmo tempo que lhes confere o direito de reivindicação junto a entidades de defesa do consumidor, nos casos de constatação de violação da padronização.

Atualmente, o cenário da padronização aparece permeado por inúmeras entidades de cunho internacional, regional e nacional, governamentais e não governamentais, que se dedicam à elaboração de padrões relativos à ergonomia de produtos de *software*. A conformidade com os padrões produzidos por estas entidades implica benefícios para os desenvolvedores (*marketing*, como um sinal de reconhecimento externo da qualidade dos produtos conformes), consumidores (direitos de reivindicação) e agentes de regulamentação (controle da qualidade dos produtos lançados no mercado de consumo).

Em nível internacional, conta-se com as iniciativas independentes e conjuntas de padronização da ISO, da IEC e da ITU, assim como a parceria estratégica destas entidades com a *World Trade Organization (WTO)*, a organização internacional que lida com regras globais de negócios entre nações, com o propósito de assegurar fluxos de exportação e importação previsíveis e livres de entraves.

Enquanto a estrutura da WTO suporta os acordos políticos necessários aos negócios internacionais, os escopos complementares, as estruturas, as competências e experiências das três principais organizações para a padronização internacional – ISO, IEC e ITU – suportam os acordos técnicos, visando o crescimento do mercado global. Informações mais detalhadas a respeito de cada uma destas organizações podem ser encontradas nos endereços indicados no Quadro 16.

Quadro 16 – Organizações Internacionais de Padronização

ORGANIZAÇÃO INTERNACIONAL	INFORMAÇÕES ONLINE
International Organization for Standardization (ISO)	www.iso.ch/
International Electrotechnical Commission (IEC)	www.iec.ch/
International Telecommunication Union (ITU)	www.itu.int/

Em nível regional, o mundo se encontra atualmente dividido em quatro blocos, cada um dos quais possui um ou mais corpos responsáveis pela harmonização técnica voluntária dos países-membros, em conjunção com corpos de padronização dos demais blocos, sob a égide das organizações responsáveis pela padronização internacional. O Quadro 17 apresenta a divisão do globo em termos da padronização regional, os corpos reconhecidos pela ISO, IEC e ITU em nível regional e os *sítes* onde se pode encontrar informações mais detalhadas a respeito de cada um deles.

Quadro 17 – Corpos de Padronização Regional reconhecidos pela ISO, IEC e ITU

REGIÃO	CORPO DE PADRONIZAÇÃO REGIONAL	INFORMAÇÃO ONLINE
AMÉRICAS	Comisión Panamericana de Normas Técnicas (COPANT)	www.copant.org/
ÁSIA E PACÍFICO	ASEAN Consultative Committee for Standards and Quality (ACCSQ)	www.aseansec.org/
	Pacific Area Standards Congress (PASC)	www.pascnet.org/default.htm
ESTADOS ÁRABES	Arab Industrial Development And Mining Organization (AIDMO)	www.arifonet.org.ma/aidmo_us/main.htm
EUROPA	Comité Européen de Normalisation (CEN)	www.cenorm.be/
	European Telecommunications Standards Institute (ETSI)	www.etsi.org/
	Comité Européen de Normalisation Electrotechnique (CENELEC)	www.cenelec.org/
	United Nations Economic Commission for Europe (UNECE)	www.unece.org/

Em nível nacional, um grande número de países possui entidades dedicadas a atividades de padronização, visando a harmonização dos interesses das sociedades, provendo-lhes referenciais através da normalização e certificação nacionais, assim como de outras atividades afins. O Anexo B apresenta uma listagem dos países que possuem representações na ISO e na IEC, discriminando o *status* de cada um deles junto a estas organizações internacionais.

Conforme pode ser observado no Anexo B, o órgão responsável pela normalização técnica que representa o Brasil junto à ISO é a *Associação Brasileira de Normas Técnicas (ABNT)*, uma entidade privada, sem fins lucrativos, fundada em 1940. A ABNT foi reconhecida como *Fórum Nacional de Normalização (ÚNICO)*, através da *Resolução Nº 07 do CONMETRO*, de 24.08.1992, sendo membro fundador da ISO, da COPANT e da *Associação Mercosul de Normalização (AMN)* [ABNT00a].

Atualmente, o *Conselho Técnico* da ABNT é composto por 47 *Comitês Brasileiros (CB)* e 2 *Organismos de Normalização Setorial (NOS)*, atuantes nas mais diversas áreas. O ABNT/CB-21 *Computadores e Processamento de Dados* é o comitê técnico de normalização ao qual está vinculado o *Subcomitê de Software (SC-21:10)*, resultante de um convênio formalizado em junho de 1992 entre a *Associação Brasileira de Normas Técnicas (ABNT)*, a *Companhia de Informática do Paraná (CELEPAR)*, a empresa *POLO de Software S/A* (Curitiba – PR) e o *Instituto de Tecnologia do Paraná (TECPAR)*. O *Subcomitê de Software (SC-21:10)* responde pelas normas de engenharia, qualidade e portabilidade de *software*, como as de implementações padronizadas relativas a linguagens, sistemas operacionais, bancos de dados e ambientes e aplicações. O subcomitê também é responsável pela coordenação da participação brasileira junto ao *ISO/IEC/JTC1/SC7 Software Engineering* e ao *ISO/IEC/JTC1/SC22 Programming Languages, their Environments and Systems Software Interfaces*.

No âmbito dos subcomitês, as *Comissões Técnicas (CT)* são responsáveis pela formação de *Comissões de Estudos (CE)*, a partir de pesquisas realizadas junto à comunidade com o propósito de verificar o interesse em estudos normativos para áreas específicas. Deste modo, as *Comissões de Estudos (CE)* instauradas passam a ser responsáveis pela elaboração de propostas de textos normativos que, mediante aprovação, darão origem a *Normas Brasileiras (NBR)*. Há duas *Comissões Técnicas (CT)* vinculadas ao *SC-21:10*, cada uma delas congregando várias *Comissões de Estudos (CE)*: a *Comissão Técnica de Engenharia de Software e Portabilidade de Software (CT-21:101)* e a *Comissão Técnica de Linguagens, Sistemas Operacionais e Bancos de Dados (CT-21:102)*, ambas formadas em 1992. O Quadro 18 discrimina as comissões de estudo formadas pelas duas comissões técnicas vinculadas ao *SC-21:10*.

Quadro 18 – *Comissões Técnicas (CT) do SC-21:10 e suas Comissões de Estudo (CE)*

SUBCOMITÊ DE SOFTWARE - SC-21:10		
COMISSÃO TÉCNICA	COMISSÕES DE ESTUDO	
CT-21:101 Comissão Técnica de Engenharia de Software e Portabilidade de Software	CE-21:101.01	Qualidade de Software
	CE-21:101.03	Gerência do Ciclo de Vida do Software
	CE-21:101.04	Avaliação de processos de Software
	CE-21:101.05	Ferramenta e Ambiente
	CE-21:101.06	Estimativa de Tamanho de Software (Ponto por Função)
	CE-21:101.07	Apoio aos Processos do Ciclo de Vida do Software
	CE-21:101.08	Ergonomia de Software
CT-21:102 Comissão Técnica de Linguagens, Sistemas Operacionais e Bancos de Dados	CE-21:102.01	Linguagem SQL
	CE-21:102.07	Linguagem JAVA
	CE-21:102.09	SQL Multimedia

A *CE-21:101.08 Ergonomia de Software* foi instituída junto ao *ABNT/CB-21/SC-21:10* em 04/1999, tendo iniciado as atividades no mês de Julho subsequente. A *CE-21:101.08 Ergonomia de Software* se reúne uma vez por mês na cidade de Curitiba para estudar o padrão internacional ISO 9241 "Ergonomic requirements for office work with Visual Display Terminals (VDTs), com o propósito de criação de uma NBR correspondente, *Requisitos ergonômicos para o trabalho de escritório com computadores*. Segundo a ABNT [ABNT00b], a *CE-21:101.08* pretende, em particular, traduzir tanto as partes introdutórias (*Partes 1 e 2*) do ISO 9241 quanto aquelas referentes às interfaces de *software (Partes 10 a 17)*. Neste sentido, a ABNT [ABNT00b] afirma que a *Parte 1 - Introdução Geral* e a *Parte 10 - Princípios de Diálogo* já foram traduzidas, aguardando somente uma revisão de estilo para que sejam encaminhadas para publicação. Por outro lado, a *Parte 11 - Orientações sobre Usabilidade* está sendo o alvo das atividades atuais da comissão.

Informações mais detalhadas sobre a estrutura e os programas de normalização técnica e certificação da ABNT podem ser obtidas via Web, no endereço www.abnt.org.br/. Informações adicionais sobre o *Subcomitê de Software (SC-21:10)* e entidades vinculadas estão disponíveis na Web no endereço www.pr.gov.br/abntsoftware/ce10108/index.html.

5.2 Aspectos da Padronização da Interação Homem-Máquina

Há menos de duas décadas, pouca atividade era registrada no âmbito das várias organizações nacionais e internacionais destinadas à padronização, no sentido de elaborarem padrões para interfaces com o usuário de sistemas computacionais. A preocupação da indústria estava direcionada para a padronização do *hardware* e, em segundo plano, para a padronização de algumas partes do *software*, e.g. linguagens ou protocolos de comunicação [Hold89, Treu94]. Portanto, pode-se afirmar que pouca atenção era prestada ao usuário.

Felizmente, esta visão tem mudado, sobretudo ao longo da última década, quando se testemunhou um aumento significativo das atividades voltadas para a padronização de interfaces com o usuário. Obviamente diversos estímulos contribuíram para a ocorrência deste fato, dentre os quais podem ser citados: (i) os resultados das pesquisas na área da interação homem-máquina e, em especial, dos processos interativos usuário-computador, descrevendo problemas com as interfaces correntes e sugerindo melhores estratégias de desenvolvimento visando o aumento da satisfação e produtividade do usuário; (ii) a conscientização dos usuários finais com a conseqüente elevação do seu grau de exigência por interfaces mais consistentes e compatíveis, bem como por dispositivos de visualização mais confortáveis; (iii) o interesse dos empregadores no aumento da produtividade de seus empregados mediante a redução de períodos de treinamento, erros e sentimentos adversos (e.g., frustração, receio, ansiedade); e (iv) a pressão das comunidades usuárias por padrões destinados à regulamentação da qualidade das interfaces dos produtos lançados no mercado de consumo [Hold93, Bill96].

Vários argumentos têm sido utilizados para fundamentar a importância da padronização voltada para a ergonomia de produtos de *software*. O primeiro deles diz respeito à focalização da atenção sobre a *consistência*, um fator relevante em interfaces usuário-computador, tanto em nível da facilidade de aprendizado quanto da facilidade de memorização dos mecanismos de interação pelo usuário, ambas facetas da usabilidade de produtos de *software* [Pree93, Stew98, Cons99]. O segundo concerne à otimização das práticas de projeto e avaliação de interfaces mediante a adoção de padrões bem concebidos, enfoque este que pode conduzir à disponibilização de produtos de *software* mais usáveis e, por conseqüente, melhores do ponto de vista do mercado de consumo, além de contribuírem para o reaproveitamento do código da interface [Bill96, Stew98]. O terceiro argumento é relativo ao aumento do conforto e bem-estar do usuário [Stew98, Cons99] e, por conseqüente, de sua satisfação.

O quarto argumento concerne à padronização como uma base para o entendimento comum entre os agentes envolvidos no projeto, manufatura, avaliação e uso de produtos [Stew98]. O quinto argumento é que a padronização estabelece uma priorização apropriada das questões relativas à interface usuário-computador no contexto maior do desenvolvimento do produto [Stew98, Cons99]. O penúltimo argumento diz respeito ao preenchimento de requisitos legais, e.g. segurança e saúde no trabalho com dispositivos de visualização⁶⁶ [Stew98]. Finalmente, o sexto e último argumento é o aumento da produtividade, promovido pela otimização da usabilidade do produto [Bill96].

⁶⁶ Este exemplo é concernente à diretiva 90/270/EEC, publicada em 1990 pela Comunidade Européia, um ato que culminou, ao final de 1992, com a transposição da referida diretiva para a legislação nacional dos países membros.

Em contrapartida, também se inserem na argumentação da importância dos padrões destinados ao projeto e avaliação de produtos de *software* ergonômicos algumas "arestas", três das quais se afiguram mais críticas. A primeira delas desponta da generalidade e do cunho de subjetividade com que a usabilidade ainda é tratada e ao reflexo deste fato sobre a consideração do atributo usabilidade como fundamento para a padronização.

Como foi evidenciado ao longo dos capítulos anteriores, em especial do segundo capítulo, a *usabilidade* de produtos é um tema de tratamento mais complexo do que a *utilidade* de produtos, visto envolver a interação de máquinas com seres humanos e, daí, todas as nuances de subjetividade que caracterizam a componente humana do processo. Apesar de tudo, a noção de usabilidade como fator de qualidade é praticamente um consenso em ergonomia [Cons99]. Um fato que reforça esta afirmativa é a existência de uma parte dedicada exclusivamente a especificações de usabilidade no padrão ergonômico internacional *ISO 9241*⁸⁷.

No segundo capítulo deste documento, foi apresentada a perspectiva segundo a qual o atributo *usabilidade* é encarado pela ISO. No modelo da ISO, a usabilidade é especificada em termos da eficiência (*efficiency*), efetividade (*effectiveness*) e satisfação (*satisfaction*) relativos à execução de tarefas por usuários. Vale a pena ressaltar que embora o escopo deste trabalho esteja restrito ao domínio das interfaces computacionais, o escopo da usabilidade transcende a interface. Além disso, o padrão *ISO 9241* como um todo se refere a requisitos ergonômicos para atividades envolvendo sistemas computacionais. O padrão *ISO 9241* trata a usabilidade como um conceito ergonômico mais genérico da qualidade, aplicável a todos os tipos de interação de um usuário com um produto em qualquer contexto de uso. O produto pode ser de *hardware* ou de *software*, o contexto pode incluir quaisquer tarefas do usuário, assim como os mais diversos fatores ambientais, sociais e físicos que podem afetar o desempenho.

A qualidade em termos de usabilidade é um aspecto cuja avaliação ocorre sobretudo para produtos de *software*, haja visto tal perspectiva ter emergido no âmbito da pesquisa de interfaces usuário-computador. É importante ter em mente que tal conceito também se estende ao *hardware* e apesar de, mais uma vez, a visão da interface tender a direcionar o foco para a usabilidade, esta incorpora a *efetividade* do desempenho, um aspecto que transcende a interface, recaindo sobre características técnicas da qualidade, e.g., funcionalidade e confiabilidade.

Dado um produto efetivo, o foco permanece direcionado para a *eficiência* do produto, o aspecto ergonômico genuíno da qualidade que é avaliado levando-se em conta os atributos da interface. Implicitamente, isto envolve (i) usuários executando tarefas com o auxílio de sistemas computacionais, mediados por uma interface e (ii) a satisfação destes usuários no tocante à mediação oferecida pela interface.

Precisamente por conta da dificuldade de uma definição mais precisa da fronteira que separa a aplicação de *software* de sua interface com o usuário, o que denota um conhecimento ainda imaturo e insuficiente sobre usabilidade, é que se questiona a padronização de interfaces usuário-computador, conforme observaram Billingsley *et al.* [Bill96].

⁸⁷ Trata-se da Parte 11 - *Especificação de Usabilidade (Usability Guidance)*, que será apresentada mais adiante, na seção 5.4.

A segunda aresta sobressai da ação inibidora dos padrões sobre a inovação em projetos de interfaces para produtos de *software*. A este respeito, Constantine e Lockwood [Cons99] comentaram que, a padronização antes da aquisição de experiência pode conduzir ao risco da adoção de padrões que, por um lado, não sejam de fato aplicáveis, devido à percepção incompleta dos princípios nos quais se fundamentam. E, por outro lado, não sejam realmente expressivos e relevantes aos interesses da organização. O que não é explicitamente dito é que, segundo esta visão, no pior destes casos os interesses da organização estarão fadados a serem lesados e, no melhor dos casos, o poder de inovação da organização poderá vir a ser limitado por uma série de normas e recomendações adotadas sem um exercício de reflexão mais aprofundado e uma ponderação mais acurada dos diversos fatores envolvidos no processo de desenvolvimento.

Entretanto, é necessário convir que tal ponto de vista reflete os interesses do desenvolvedor, nem sempre consonantes com os interesses do usuário, conforme ressaltou Engel [Enge99], ao afirmar que os padrões que conferem, de algum modo, poder ao consumidor nem sempre são do interesse dos desenvolvedores. Segundo Engel [Enge99], a necessidade da padronização orientada ao *software*, que dite como o usuário interagirá com diferentes produtos, é menos óbvia do que aquela orientada ao *hardware*, tornando bastante comuns as "evasões" aos padrões no âmbito das interfaces dos produtos de *software*.

Engel [Enge99] também ressaltou que muitos desenvolvedores freqüentemente costumam apregoar "extensões funcionais" em seus produtos, na verdade funcionalidades adicionais de curto prazo que ocultam "armadilhas" na implementação e limitações na aplicabilidade, na maioria dos casos constatadas somente a médio e longo prazos pelos consumidores. A aquisição de produtos conformes a padrões conferem direitos ao usuário no tocante à exigência do cumprimento de características apregoadas, além de favorecerem o aumento da competitividade entre desenvolvedores, impulsionando a inovação como um dos modos de diferenciação.

A terceira aresta que desponta na discussão sobre a importância da padronização centrada na usabilidade diz respeito à postura expressa pelos desenvolvedores que acreditam que os testes são o único meio de garantir a usabilidade de produtos.

De acordo com a discussão anterior, é ilusório crer que a adoção de recomendações contidas em padrões é suficiente para garantir o sucesso do processo de desenvolvimento de produtos e/ou a usabilidade dos produtos finais. Como foi enfatizado, os padrões referentes a aspectos do processo de desenvolvimento e, em especial, dos processos interativos usuário-computador ainda estão em um estágio emergente, tendo em vista muitos deles ainda se encontrarem em estágios de elaboração, discussão e votação.

Além disso, o uso dos documentos finais dos padrões requer cautela e ponderação na interpretação, pois o que oferecem, via de regra, são conjuntos de listas de inspeção (*checklists*) para que equipes de projeto e avaliação do mundo inteiro sejam encorajadas a centrarem seus processos e produtos no usuário. Redobre-se a cautela e a ponderação no que diz respeito ao uso de versões preliminares de padrões (e.g., *working draft*, *draft international standard*, *final draft international standard*), tendo em vista ainda se encontrarem em fase de discussão e/ou votação e, por conseguinte, sujeitas a modificações.

No tocante ao conteúdo, os padrões destinados aos produtos de *software* costumam ser enquadrados em duas grandes categorias, a saber: (i) *padrões para processos*, concernentes essencialmente a diferentes aspectos de projeto e desenvolvimento de *software*, em especial à especificação de procedimentos e processos a serem adotados; e (ii) *padrões para produtos*, relativos à ergonomia e questões de acessibilidade de produtos, em especial à especificação de atributos da interface com o usuário⁶⁸ [INUS97, Crer98].

No que diz respeito aos componentes do sistema computacional, Stewart [Stew98, Stew00] ainda apontou um segundo modo de subdivisão: (i) padrões que tratam da ergonomia do *hardware*, i.e., de como o suporte computacional físico se adequa à componente humana do processo interativo; e (ii) padrões que tratam da ergonomia do *software*, i.e., de como o suporte lógico do sistema e, em especial, da interface com a componente humana se adequam às características dos usuários, às tarefas que estes objetivam executar e aos contextos de uso.

5.3 Padrões Destinados à Interação Homem-Máquina

Antes de apresentar os padrões internacionais ISO e ISO/IEC relativos a diferentes aspectos da interação homem-máquina, vale a pena descrever a *Classificação Internacional para Padrões*⁶⁹, que consiste de três níveis hierárquicos: *área (nível 1)*, *grupo (nível 2)* e *sub-grupo (nível 3)*.

O *nível 1* concerne às áreas das atividades de padronização, de um modo mais abrangente. Cada *área* corresponde a uma notação de dois dígitos, e.g., **13 ENVIRONMENT. HEALTH PROTECTION. SAFETY**. As áreas são subdivididas em *grupos (nível 2)*, cuja notação consiste da notação da área acrescida de um ponto e um número composto de três dígitos, e.g., **13.180 Ergonomics**. Diversos grupos são divididos em *sub-grupos (nível 3)*, que não é o caso do grupo **13.180 Ergonomics**, o qual não possui sub-grupos, mas é o caso do grupo **17.040 Linear and angular measurements**, o qual possui quatro sub-grupos, o primeiro dos quais é o **17.040.01 Linear and angular measurements in general**.

Vários grupos e sub-grupos são seguidos por notas de escopo e/ou de referência, impressas em itálico, e.g.

- **35.240.80 IT applications in health care technology**
Including computer tomography

O mesmo padrão pode ser classificado em vários grupos e/ou sub-grupos, e.g. o padrão ISO 9241 está classificado como **13.180 Ergonomics** e como **35.180 IT terminal and other peripheral equipment**. A classificação é atualizada conforme as necessidades. Qualquer usuário pode submeter propostas para modificações e/ou adendos ao ICS. Para tanto, o usuário deverá encaminhar a proposta para o *ISO Central Secretariat*, que a repassará para a *ICS Maintenance Agency*, a fim de que seja apreciada. Outro tópico que merece esclarecimento refere-se ao *número de referência (reference number)* das publicações da ISO, que consiste de um prefixo, um número

⁶⁸ Alguns padrões para produtos apresentam a especificação de requisitos em nível do desempenho e não de atributos do produto, descrevendo os usuários, tarefas e contextos de uso e analisando a usabilidade do produto em termos dos níveis de desempenho e satisfação do usuário a serem atingidos.

⁶⁹ *International Classification for Standards (ICS)*

de série e o ano da publicação. O prefixo será usualmente ISO, indicando que a publicação é um padrão internacional da ISO. Entretanto, é possível encontrar também os prefixos ISO/IEC ou ISO/CIE, a fim de indicar que se trata de uma publicação conjunta da ISO e IEC (usualmente desenvolvida pelo *Joint ISO/IEC Technical Committee JTC1*) ou da ISO e CIE (*Commission Internationale de l'Éclairage*).

O prefixo também pode conter uma indicação do tipo do documento publicado, conforme sintetizado no Quadro 19 e, caso contrário, é sucedido por um número de série que pode incluir um número referente à parte, que aparece separado por um hífen do número principal. O número de série, por sua vez, é seguido pelo ano da publicação, separado do número de série por um sinal de dois pontos (:).

Quadro 19 – Prefixos indicadores de tipos de publicações da ISO.

PUBLICAÇÕES DA ISO	
PREFIXO	DESCRIÇÃO
Amd	Emenda (<i>Amendment</i>) - documento normativo destinado à alteração de elementos técnicos normativos de outro documento, sendo desenvolvido consensualmente e aprovado conforme os procedimentos relevantes ao documento a ser emendado.
Cor	Corrigenda Técnica (<i>Technical Corrigendum</i>) - documento destinado à correção de erros técnicos ou ambigüidades contidos em um documento normativo ou à correção de informações obsoletas, visto que a modificação não exerce efeitos sobre os elementos técnicos normativos do documento corrigido.
Guide	Guia (<i>Guide</i>) - documento tratando de questões não normativas relacionadas à padronização internacional (e.g., a aplicação de padrões "horizontais") ou documento normativo desenvolvido por um comitê diferente da estrutura TC/SC (e.g., comitê de desenvolvimento de planos de ação).
ISP	Perfil Internacional Padronizado (<i>International Standardized Profile</i>) - documento harmonizado e consonante internamente, que identifica um padrão ou um grupo de padrões, juntamente com opções e parâmetros, necessários à execução de uma função ou de uma sequência de funções.
ITA	Acordo Industrial Técnico (<i>Industry Technical Agreement</i>) - documento técnico desenvolvido a partir de uma reunião de trabalho fora das estruturas técnicas da ISO e IEC.
PAS	Especificação Disponível ao Público (<i>Publicly Available Specification</i>) - documento técnico, representante de um consenso entre especialistas, aprovado por uma maioria simples de membros P (participantes) de um TC ou SC e publicado para atender a uma necessidade urgente de mercado. Uma PAS não pode ser conflitante com um padrão internacional existente, embora sejam permitidas PAS concorrentes sobre o mesmo tema.
R	Recomendação (<i>Recommendation</i>) - designação usada até 1972, quando a ISO começou a publicar padrões internacionais. À medida em que vêm sendo revisadas, as recomendações da ISO têm-se convertido gradualmente em padrões internacionais. Todavia, um número limitado de recomendações da ISO ainda permanece válido e disponível.
TR	Relatório Técnico (<i>Technical Report</i>) - documento informativo contendo informações de natureza distinta (e.g., coletânea de dados) daquela normalmente publicada em um padrão internacional.
TTA	Avaliação de Tendências Tecnológicas (<i>Technology Trends Assessment</i>) - documento publicado com o propósito de atender a carência de colaboração global em questões de padronização durante os estágios preliminares de inovações técnicas, delineado o estado da arte ou as tendências em áreas emergentes. Resulta tipicamente de iniciativas de trabalho e pesquisa que precedem a padronização.
TS	Especificação Técnica (<i>Technical Specification</i>) - documento normativo, desenvolvido segundo procedimentos consensuais e aprovado por dois terços dos membros P do comitê responsável, que é publicado quando não há suporte para a publicação de um padrão internacional, quando o tema central ainda se encontra em desenvolvimento técnico ou em casos em que há a possibilidade futura, mas não imediata, de harmonização para um padrão internacional. Obedece às mesmas regras das PAS.

Além dos prefixos indicadores apresentados no Quadro 19, os números de referência das publicações da ISO podem conter abreviações que indicam o estágio de evolução do documento. O Quadro 20 apresenta as abreviações adicionadas aos números de referência das publicações ISO, com o propósito de indicar os diferentes estágios de elaboração de documentos vigentes na estrutura da ISO. Por sua vez, o Quadro 21 detalha a codificação adotada pela ISO para os diferentes estágios do processo de elaboração de um padrão internacional apresentada resumidamente no Quadro 20, enquanto o Quadro 22 apresenta a descrição das principais abreviações, apenas listadas no Quadro 20.

Quadro 20 – Abreviações adotadas durante os estágios de desenvolvimento de publicações ISO

Estágio	Abreviação	Especificação	
20 Preparatório	20.00	AWI	Item de Trabalho Aprovado, ainda sem versão preliminar de trabalho disponível
		AWI Amd	Proposta Aprovada para Emenda
		AWI TR ou TS	Proposta Aprovada para Relatório Técnico ou Especificação Técnica
	20.20	WD	Esboço Preliminar de Trabalho
		WD Amd	Esboço Preliminar de Trabalho de Emenda
		WD TR ou TS	Esboço Preliminar de Trabalho de Relatório Técnico ou Especificação Técnica
30 Comitê	30.00 a 30.99	CD	Esboço Preliminar de Comitê
		CD Amd	Esboço Preliminar de Comitê de Emenda
		CD Cor	Esboço Preliminar de Comitê de Corrigenda Técnica
		CD TR ou TS	Esboço Preliminar de Comitê de Relatório Técnico ou Especificação Técnica
		DTR	Esboço Preliminar de Relatório Técnico
		PD Amd	Esboço Preliminar de Emenda Proposta
40 Consulta	40.00 a 40.99	DIS	Esboço Preliminar de Padrão Internacional
		DAmd	Esboço Preliminar de Emenda
		FCD	Esboço Final de Comitê
		FPDISP	Esboço Final proposto para Perfil Internacional Padronizado
50 Aprovação	50.00 a 50.99	FDIS	Esboço Final de Padrão Internacional
		FD Amd	Esboço Final de Emenda
		PRF	Prova de novo Padrão Internacional
		PRF Amd	Prova de Emenda
		PRF TR ou TS	Prova de de Relatório Técnico ou Especificação Técnica
		PRF Suppl	Prova de Suplemento
60 Publicação	60.00 a 60.99	ISO	Padrão Internacional
		ISO/TR ou TS	Relatório Técnico ou Especificação Técnica
		Amd	Emenda
		Cor	Corrigenda Técnica

As informações contidas no Quadro 21 são importantes tanto quando se pretende adquirir um determinado padrão de interesse quanto em situações nas quais se deseja conhecer o *status* corrente de um padrão em versão preliminar (*draft*) e acompanhar seu progresso ou verificar se uma versão final se encontra sendo revisada, se está sendo retirada de circulação ou se já o foi.

Quadro 21 – Detalhamento da codificação dos estágios de um documento ISO

SUB-ESTÁGIO							
ESTÁGIO	00	20	60	90			
	Registro	Início da Ação Principal	Conclusão da Ação Principal	Decisão			
				92	93	98	99
				Repetição de uma etapa anterior	Repetição da etapa corrente	Abandono	Confirmação
00 Preliminar	00.00 Recebimento da proposta de novo projeto	00.20 Revisão da proposta de novo projeto	00.60 Circulação do sumário da revisão			00.98 Descarte da proposta de novo projeto	00.99 Aprovação da votação da proposta de novo projeto
10 Proposição	10.00 Registro da proposta de novo projeto	10.20 Início da votação de novo projeto	10.60 Circulação do sumário da votação	10.92 Devolução da proposta ao proponente, para definição ulterior		10.98 Rejeição do novo projeto proposto	10.99 Aprovação do novo projeto proposto
20 Preparação	20.00 Registro do novo projeto no programa de um TC/SC	20.20 Início do estudo da WD	20.60 Circulação do sumário dos comentários			20.98 Cancelamento do projeto	20.99 Aprovação do registro da WD como CD
30 Comitê	30.00 Registro da CD	30.20 Início do estudo e votação da CD	30.60 Circulação do sumário dos comentários/votação	30.92 Devolução da CD ao WIG associado		30.98 Cancelamento do projeto	30.99 Aprovação do registro da CD como DIS
40 Consulta	40.00 Registro da DIS	40.20 Início da votação da DIS: 5 meses	40.60 Despacho do sumário da votação	40.92 Circulação do relatório geral : devolução da DIS ao TC ou SC	40.93 Circulação do relatório geral : decisão da nova votação da DIS	40.98 Cancelamento do projeto	40.99 Circulação do relatório geral : Aprovação do registro da DIS como FDIS
50 Aprovação	50.00 Registro da FDIS para aprovação formal	50.20 Início da votação da FDIS: 2 meses. Cópia enviada ao secretariado	50.60 Despacho do sumário da votação. Cópia devolvida pelo secretariado	50.92 Devolução da FDIS ao TC ou SC associado		50.98 Cancelamento do projeto	50.99 Aprovação da publicação da FDIS
60 Publicação	60.00 IS no prelo		60.60 IS publicado				
90 Revisão		90.20 IS em revisão periódica	90.60 Despacho do sumário da revisão	90.92 IS a ser revisado	90.93 IS revisado		90.99 Proposição de exclusão do IS por TC ou SC
95 Exclusão		95.20 Início da votação para exclusão do IS	95.60 Despacho do sumário da votação	95.92 Decisão de não exclusão do IS			95.99 Exclusão do IS

Quadro 22 – Abreviações adotadas pela ISO ao longo dos diferentes estágios de elaboração dos documentos de padrões internacionais

ABREVIÇÃO	DESCRIÇÃO DETALHADA
WI	Item de Trabalho (<i>Work Item</i>) - tópico aprovado e reconhecido por um grupo de trabalho, a ser enviado àqueles que lidam com um ou mais padrões publicados.
WD	Esboço Preliminar de Trabalho (<i>Working Draft</i>) - primeira versão, parcial ou completa, do documento de um padrão proposto.
CD	Esboço Preliminar de Comitê (<i>Committee Draft</i>) - documento posto em circulação no âmbito do comitê de trabalho e dos comitês nacionais correspondentes. Votação e aprovação necessárias para que o documento evolua para o próximo estágio.
DIS	Esboço Preliminar de Padrão Internacional (<i>Draft International Standard</i>) - versão preliminar do padrão, posta amplamente em circulação, para comentários públicos por organizações de elaboração de padrões nacionais. Votação e aprovação necessárias para que o documento evolua para o próximo estágio.
FDIS	Esboço Final de Padrão Internacional (<i>Final Draft International Standard</i>) - versão preliminar final do padrão, posta em circulação para votação antes da adoção como um padrão internacional.
IS	Padrão Internacional (<i>International Standard</i>) - versão final do documento, posta à venda pela ISO.
EN	Norma Européia (<i>European Norm</i>) - abreviação acrescentada ao número de referência de um documento que também é adotado como norma pela Comunidade Européia (e.g. IS/EN).

A Fig. 11 exemplifica a interpretação dos números de referência de duas publicações da ISO - um Esboço Preliminar de Especificação Técnica e uma Parte de Padrão Internacional.

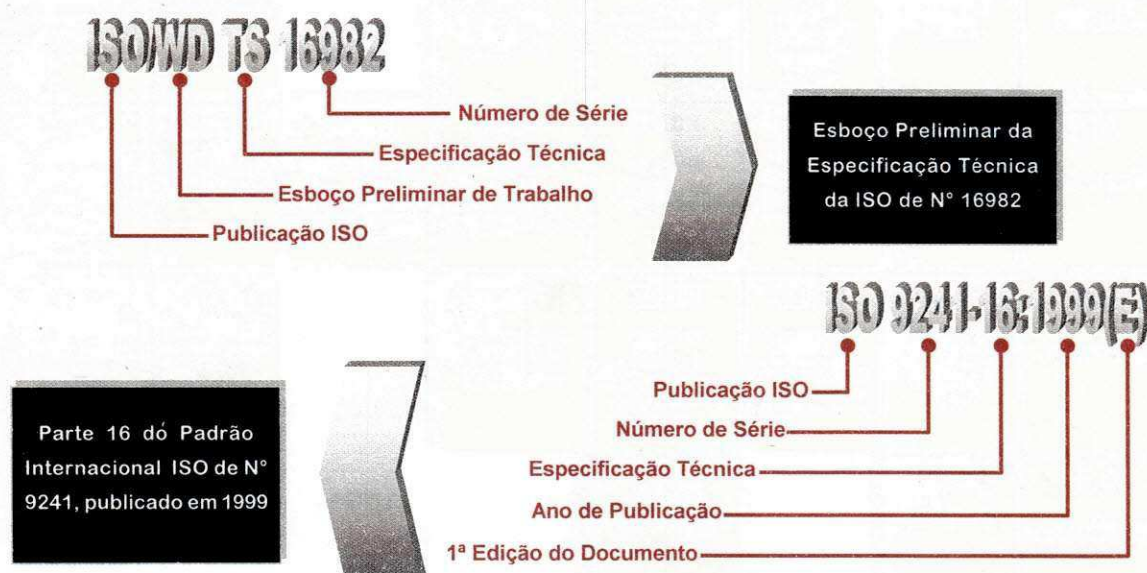


Fig. 11 – Exemplos de interpretação de números de referência de documentos ISO.

O Anexo C lista as publicações atuais desenvolvidas pela ISO, algumas das quais em conjunto com a IEC, relativas a diferentes aspectos da interação homem-máquina, destacando comitês e sub-comitês responsáveis, *status* dos documentos, datas das atribuições dos *status* e codificação referente à edição, número de páginas e preço (entre parênteses e em itálico, após a especificação). O padrão que se insere de fato no contexto desta pesquisa é o ISO 9241, motivo pelo qual suas partes serão apresentadas separadamente mais adiante.

Nos EUA, a *Human Factors and Ergonomics Society (HFES)* foi aprovada como uma organização de desenvolvimento de padrões (*Standards Developing Organization - SDO*) no que diz respeito à elaboração de um padrão nacional americano para interfaces usuário-computador. O referido padrão se encontra atualmente em desenvolvimento pelo *HFES/HCI 200 Committee*, sob os auspícios do *HFES Technical Standards Committee*. Quando o documento estiver completo, o esboço do padrão passará por um estágio de revisão pública, conhecido como *ANSI Canvass Process*. Participam deste processo de aprovação quaisquer indivíduos interessados e envolvidos com o contexto da padronização, através da revisão do documento e da formulação de comentários diretamente para a SDO. Esta, por sua vez, revisará o esboço, a fim de executar tantas reformulações quantas forem tecnicamente apropriadas, visando atender as reivindicações dos participantes do processo de revisão pública.

O *HFES 200*, como foi denominado o padrão em desenvolvimento, pretende apresentar requisitos e recomendações para interfaces de *software*, com enfoque central nas aplicações computacionais de escritório e pessoais, relacionando-se, neste sentido, com o padrão multi-partes *ISO 9241*, que será discutido na seção seguinte. Segundo a IBM [IBM99], este padrão ANSI/HFES adotará, modificará ou estenderá o conteúdo das diferentes partes do *ISO 9241*, nos casos em que se fizer necessário, a fim de abranger as mais recentes tecnologias e pesquisas e refletir mais fielmente a perspectiva americana. Assim, o *HFES 200* será compatível, mas não igual, à série de Partes 10 a 17 do padrão internacional *ISO 9241*, englobando todas as técnicas de diálogo tratadas nos padrões da ISO e oferecendo adicionalmente recomendações e orientação no tocante às questões de entrada e saída de voz, uso de cor e acessibilidade.

A seção referente ao uso de cor, que conterá diretrizes e recomendações para o uso apropriado de recursos cromáticos em interfaces usuário-computador, visando a otimização da eficiência do usuário na execução de tarefas, pretende complementar e estender o conteúdo das Partes 3, 8 e 12 do *ISO 9241*, assim como do padrão ANSI/HFES 100-199X.

No Brasil, conforme anteriormente comentado, a única iniciativa vigente no sentido da padronização de aspectos da interação usuário-computador são os estudos da *CE-21:101.08*, voltados essencialmente para a tradução das partes introdutórias (Partes 1 e 2) e daquelas referentes às interfaces de *software* (Partes 10 a 17) do padrão internacional *ISO 9241*. De acordo com as informações da ABNT [ABNT00b], a *Parte 1 - Introdução Geral* e a *Parte 10 - Princípios de Diálogo* já foram traduzidas, aguardando apenas uma revisão de estilo a fim de poderem ser publicadas, enquanto a *Parte 11 - Orientações sobre Usabilidade* é o foco atual de tradução pela comissão *CE-21:101.08*.

5.4 O Padrão ISO 9241

A preocupação com a ergonomia de terminais de visualização cresceu no final da década de 70, sobretudo no tocante à possível fadiga visual e deterioração da visão dos usuários devido ao uso prolongado de terminais de vídeo, especialmente aqueles com baixa qualidade de apresentação visual. Entretanto, é importante chamar a atenção que vários estudos realizados nos últimos anos têm mostrado o envelhecimento humano como o principal causador da deterioração visual. Muitos indivíduos só constatarem tal deterioração através do desconforto devido ao uso intensivo de terminais de vídeo. Uma vez que o uso, mesmo não intensivo, desses dispositivos pode ser

visualmente fatigante, pode-se incorrer no erro de atribuir a necessidade de corretivos visuais ao uso excessivo de terminais de vídeo.

Foram questões relativas ao uso de dispositivos de visualização que motivaram a proposição de um novo tópico de trabalho. Assim, sob os auspícios do comitê de *Tecnologia da Informação*, foi criado o *ISO/TC 159*, o comitê de ergonomia da ISO. O tópico de trabalho foi atribuído ao sub-comitê *ISO/TC 159/SC4 Sinais e Controle*, tendo a reunião inaugural ocorrido no BSI (Manchester – UK), em 1983, com a participação de delegações de diversos países e a tomada de algumas decisões-chaves.

Na época, a proliferação sistemas computacionais para escritórios (*office based systems*) contribuiu para a focalização dos interesses em tarefas de escritórios (e.g., processamento de texto, uso de planilhas eletrônicas) ao invés de iniciativas de inclusão de CAD ou aplicações de controle de processos, conforme relatou Stewart [Stew00], membro do comitê.

Na ocasião, também foi decidida a elaboração de um padrão composto por uma série de partes, visando cobrir a extensa gama de tópicos de ergonomia que o comitê acreditava precisar de atenção no sentido da otimização da ergonomia do trabalho envolvendo terminais de vídeo. De início, foram identificados seis tópicos e definidos grupos de trabalho responsáveis por cada um deles, conforme ilustrado no Quadro 23. Passaram-se quase sete anos antes da publicação das primeiras partes do *ISO 9241* [Stew00], um padrão constituído de dezessete partes que levou mais de quinze anos para ser concluído [Blan00]. A primeira parte publicada foi a *Parte 2* (15/07/1992), cerca de uma década após a reunião inaugural do sub-comitê *ISO/TC 159/SC4 Sinais e Controle* enquanto a primeira edição da última parte, a *Part 9 Requirements for non-keyboard input devices*, datada de 15/02/2000.

Quadro 23 – Grupos de Trabalho do TC159/SC4 da ISO

GRUPO DE TRABALHO	TÓPICO CORRESPONDENTE
WG1	Fundamentos de métodos de controle e sinalização
WG2	Especificações de dispositivos de visualização
WG3	Especificações ambientais, de locais de trabalho e de controle
WG4	Especificações de tarefas (desativado)
WG5	Ergonomia de software e diálogo usuário-computador
WG6	Processos de projeto centrados no usuário para sistemas interativos
WG8	Projeto ergonômico de centros de controle

O Quadro 24 discrimina as dezessete partes do padrão *ISO 9241 - Ergonomic requirements for office work with visual display terminals (VDTs)*⁷⁰ apresentando o *status* atual de cada uma delas, inclusive das duas emendas elaboradas para as *Partes 1 e 3*. Aparecem em destaque as *Partes 11, 14, 15 e 16*, em virtude de haverem sido utilizadas no processo de inspeção conduzido no âmbito desta pesquisa, como será visto mais detalhadamente na última seção deste capítulo.

⁷⁰ Especificações ergonômicas para o trabalho de escritório com terminais de vídeo.

Quadro 24 – ISO 9241 – Partes, Grupos de Trabalho e Status Atual

ISO 9241			
Ergonomics requirements for office work with visual display terminals (VDTs)⁷¹			
Parte	Denominação	Grupo de Trabalho Responsável	Status
1:1997	General Introduction	WG6	IS/EN
2:1992	Guidance on task requirements	WG4 (extinto)	IS/EN
3:1992	Visual display requirements	WG2	IS/EN
4:1998	Keyboard requirements	WG3	IS/EN
5:1998	Workstation layout and postural requirements	WG3	IS/EN
6:1999	Environmental requirements	WG3	IS/EN
7:1998	Display requirements with reflections	WG2	IS/EN
8:1997	Requirements for displayed colours	WG2	IS/EN
9:2000	Requirements for non-keyboard input devices	WG3	IS/EN
10:1996	Dialogue principles	WG5	IS/EN
11:1998	Guidance on usability specification and measures	WG5	IS/EN
12:1998	Presentation of information	WG5	IS/EN
13:1998	User guidance	WG5	IS/EN
14:1997	Menu dialogues	WG5	IS/EN
15:1997	Command dialogues	WG5	IS/EN
16:1999	Direct manipulation dialogues	WG5	IS/EN
17:1998	Form filling dialogues	WG5	IS/EN
3:1992/FD Amd 1	Visual perception test	WG2	FDIS
1:1997/FD Amd 1	Description and application of the software parts	WG6	FDIS

Nota: As partes no estágio IS aparecem em **negrito**

Segundo Stewart [Stew00], a estrutura do padrão foi definida em uma época em que as distinções entre *hardware* e *software* eram bem mais nítidas e a visualização através de tubos de raios de catódicos monocromáticos (CRT) eram a norma. O comitê técnico responsável decidiu dedicar as seis primeiras partes do padrão ISO 9241 eminentemente ao *hardware*, acreditando que se tratava de um tópico mais fácil e menos consencioso de padronizar, muito embora também reconhecesse a importância do projeto do trabalho e da tarefa no contexto do *hardware*, de modo que inseriu a *Parte 2* antes das partes mais técnicas orientadas ao *hardware*, com o propósito de favorecer a segurança, o conforto e a eficiência dos usuários durante a execução de suas tarefas.

Eis porque as 9 primeiras partes do ISO 9241 destacam-se claramente das demais, por tratarem de aspectos do *hardware* mais diretamente envolvido com os processos interativos usuário-computador (dispositivos de entrada e saída da informação) ou de condições ambientais (visuais, acústicas e térmicas) propiciadoras do conforto, da segurança e da produtividade do usuário durante a execução de suas tarefas. O Quadro 25a sumaria o conteúdo de cada uma das nove partes iniciais do padrão ISO 9241, sugerindo os segmentos da comunidade usuária de padrões⁷² mais favorecidos com as recomendações contidas em cada parte.

⁷¹ Requisitos ergonômicos para o trabalho de escritório com terminais de visualização. Nesta tradução, o trabalho com terminais de visualização sugere o uso de sistemas computacionais, visão corroborada pela CE-21:101.08 da ABNT, cujo objetivo é o estudo do padrão ISO 9241 com o propósito de adaptação do conteúdo para uma NBR correspondente, intitulada *Requisitos ergonômicos para o trabalho de escritório com computadores*.

⁷² Fabricantes, projetistas e avaliadores de interfaces e profissionais envolvidos com a licitação e certificação de produtos.

Quadro 25a – ISO 9241 – Síntese do conteúdo das Partes 1 a 9 e alvos correlatos.

ISO 9241		
Ergonomics requirements for office work with visual display terminals (VDTs) – Partes 1 a 9		
	Conteúdo	Comunidade-ANV
1	Apresentação das partes do padrão ISO 9241 e explanação de alguns princípios básicos concernentes. Recomendações sobre o uso do padrão e descrição de como a análise de conformidade com o ISO 9241 deverá ser relatada.	Público em geral envolvido com processos interativos usuário-computador que usam terminais de visualização.
2	Orientação sobre a identificação e especificação de requisitos de tarefas envolvendo terminais de visualização. Indicação de como tais requisitos podem ser incorporados ao projeto do sistema e ao processo de implementação.	Projetistas e avaliadores de tarefas e atividades envolvendo o uso de terminais de visualização.
3	Especificação de requisitos de projeto de <i>hardware</i> de visualização usado em escritórios (e, por extensão, de terminais usados em aplicações de propósito geral), visando o acesso confortável, seguro e eficiente à informação ao longo da tarefa. Apresentação de uma proposição para o teste de desempenho do usuário.	Fabricantes e profissionais atuantes na licitação ou certificação de terminais de vídeo, usuários finais.
4	Especificação de características ergonômicas de teclado alfanumérico e especificações de projeto de sistemas de escritório envolvendo teclados alfanuméricos. Enquanto a ISO 9241-4:1998 trata de aspectos ergonômicos, as 8 partes do ISO/IEC 9995:1994 - <i>Information Processing - Keyboard Layouts for Text and Office Systems</i> tratam dos diferentes aspectos de layouts de teclados.	Projetistas e avaliadores de teclados e dispositivos de entrada baseados em teclados, fabricantes e profissionais atuantes na licitação ou certificação de teclados, usuários finais.
5	Especificação de requisitos ergonômicos para estações de trabalho com terminais de visualização, a fim de que o usuário adote uma postura confortável e eficiente.	Projetistas e avaliadores de ambientes de trabalho para usuários de estações de trabalho com terminais de visualização.
6	Especificação de requisitos ergonômicos para ambientes de trabalho envolvendo o uso de terminais de visualização, visando oferecer ao usuário condições ambientais (visuais, acústicas e térmicas) confortáveis, seguras e produtivas.	Projetistas e avaliadores de ambientes de trabalho para usuários de estações de trabalho com terminais de visualização.
7	Especificação e detalhamento de métodos de mensuração da ofuscação e de reflexos produzidos na superfície de telas de dispositivos de visualização, inclusive aquelas com superfícies anti-reflexivas.	Fabricantes de monitores visando tratamentos anti-reflexivos que não degradem a qualidade da imagem.
8	Especificação de requisitos ergonômicos para terminais de visualização em cores, suplementando os requisitos monocromáticos da Parte 3.	Fabricantes e profissionais atuantes na licitação ou certificação de terminais de vídeo, usuários finais.
9	Especificação de requisitos ergonômicos para dispositivos de entrada apontadores (e.g., <i>mouse</i> , <i>trackball</i>), possíveis de ser empregados em associação com um terminal de vídeo. Inclusão de um procedimento de teste de desempenho.	Fabricantes, projetistas, avaliadores e profissionais atuantes na licitação ou certificação de dispositivos de entrada diferentes de teclados, usuários finais.

Posteriormente, as partes relativas ao *software* (Partes 10 a 17) foram acrescentadas, buscando cobrir os diferentes estilos de interação usuário-sistema possíveis na ocasião. Também foram adicionadas outras partes concernentes ao *hardware*, relativas a reflexões (Parte 7), terminais coloridos (Parte 8) e outros dispositivos de entrada além do teclado (Parte 9). Stewart [Stew00] comentou, na qualidade de integrante do comitê técnico para a elaboração do ISO 9241,

que a estrutura do padrão não é muito centrada no usuário, em virtude de ser um reflexo das questões práticas e dos avanços vertiginosos no *hardware* e *software* computacional ocorridos desde a concepção das diferentes partes do padrão. O Quadro 25b sumaria o conteúdo de cada uma das oito últimas partes do padrão *ISO 9241*, complementando a concepção do Quadro 25a.

Quadro 25b – *ISO 9241* – Síntese do conteúdo das Partes 10 a 17 e alvos correlatos.

ISO 9241		
Ergonomics requirements for office work with visual display terminals (VDTs) – Partes 10 a 17		
Parte	Conteúdo	Comunidade-Alvo
10	Apresentação de princípios ergonômicos de alto nível aplicáveis ao projeto de diálogos usuário-computador, dentre os quais adequação à tarefa, ao aprendizado e à individualização, conformidade com as expectativas do usuário, auto-descrição, controle de ações e tolerância a erros, a partir de cenários que indicam prioridades relativas e importância dos vários princípios em aplicações práticas.	Projetistas e avaliadores de diálogos usuário-sistema e navegação através de interfaces.
11	Orientação sobre especificação ou avaliação da usabilidade em termos de medidas de desempenho e satisfação do usuário. Descrições explícitas do contexto de uso do produto (<i>hardware</i> , <i>software</i> ou serviço), procedimentos de avaliação a serem conduzidos e critérios a serem adotados nas medidas de usabilidade de produtos. Explicação de como a usabilidade de um produto pode ser especificada e avaliada como parte de um sistema de qualidade, e.g., conforme com a ISO 9001.	Fabricantes de produtos, projetistas e avaliadores de interfaces usuário-computador, profissionais atuantes em licitação e/ou certificação de produtos, usuários finais.
12	Apresentação de aspectos ergonômicos específicos envolvidos na representação e apresentação das informações em terminais de visualização, incluindo orientação sobre formas de representação de informações complexas a partir de códigos alfanuméricos e gráficos/simbólicos, <i>layout</i> e projeto de telas e uso de janelas. Condensação do material mais útil e relevante do acervo substancial de referências disponíveis na literatura sobre projeto de interfaces.	Projetistas e avaliadores de interfaces usuário-sistema baseadas em terminais de visualização.
13	Apresentação de recomendações para o projeto e a avaliação de atributos de orientação do usuário de interfaces de <i>software</i> , incluindo <i>Prompts</i> , Retorno da Informação, <i>Status</i> , <i>Ajuda Online</i> e Manipulação de Erros.	Profissionais envolvidos com o treinamento do usuário e/ou suporte ao usuário, usuários finais.
14	Apresentação de recomendações para o projeto ergonômico de diálogos via menus, englobando estrutura de menus, navegação, seleção e execução de opções e apresentação de menus por várias técnicas, e.g. janelas, painéis, botões, campos.	Projetistas e avaliadores de interfaces usuário-computador com interação via menus, usuários finais.
15	Apresentação de recomendações para o projeto ergonômico de diálogos via linguagens de comando, englobando estrutura, sintaxe e representação de comandos, considerações de entrada e saída, retorno da informação e ajuda.	Projetistas e avaliadores de interfaces usuário-computador com interação via linguagem de comandos, usuários finais.
16	Apresentação de recomendações para o projeto ergonômico de diálogos via manipulação direta, incluindo a manipulação de objetos e o projeto de metáforas, objetos e atributos e GUI com recursos de manipulação direta.	Projetistas e avaliadores de interfaces usuário-computador com interação via manipulação direta, usuários finais.
17	Apresentação de recomendações para o projeto ergonômico de diálogos via preenchimento de formulários, cobrindo considerações estruturais, de entrada e saída e navegação em formulários.	Projetistas e avaliadores de interfaces usuário-computador com interação via preenchimento de formulários.

Stewart [Stew00] ainda ressaltou que o comitê antecipou um conjunto de considerações sobre atividades de projeto da interação usuário-computador e sobre como o padrão poderia lhes dar suporte. Dentre as atividades incluídas nesse conjunto, o autor destacou: (i) a análise e a definição de especificações do sistema; (ii) o projeto do diálogo usuário-sistema e a navegação através da interface; (iii) o projeto ou a seleção de modos de visualização da informação; (iv) o projeto ou a seleção de mecanismos de entrada via teclados e outros dispositivos de entrada; (v) o suporte e o treinamento aos usuários; e (vi) o projeto de tarefas e atividades do usuário.

Acrescente-se às atividades discriminadas por Stewart [Stew00] as ações relativas à avaliação de interfaces usuário-computador, complementares às atividades de projeto e passíveis de serem conduzidas ao longo de todo o ciclo de vida do produto. Estas ações avaliatórias, realizadas tanto em nível do *hardware* mais diretamente associado ao processo interativo do usuário com o sistema quanto em termos da componente de *software* da interface mediadora da comunicação entre o usuário e a aplicação, também podem ser auxiliadas pelas orientações oferecidas pelas diferentes partes do padrão ISO 9241.

A última seção deste capítulo descreverá o procedimento metodológico empregado na análise de conformidade do produto selecionado para avaliação no âmbito desta pesquisa, o *MATLAB v. 5.3 (The MathWorks Inc.)*.

5.5 Metodologia Adotada na Inspeção do Produto-Alvo com Base nas Partes 11, 14, 15 e 16 do Padrão ISO 9241

Conforme discutido na seção anterior, as nove primeiras partes do padrão distinguem-se nitidamente das demais, pois enquanto as oito últimas partes referem-se predominantemente a diferentes aspectos ergonômicos em nível do *software*, as Partes 1 a 9 tratam eminentemente de diferentes aspectos do *hardware* mais diretamente relacionado com a interação usuário-computador ou de condições ambientais (visuais, acústicas e térmicas) concernentes ao conforto, segurança e produtividade do usuário durante a realização de suas atividades.

Quadro 26 – ISO 9241 – Essência das partes constituintes.

ISO 9241	
Ergonomics requirements for office work with visual display terminals (VDTs)	
Part	
1:1997	Considerações introdutórias
2:1992	Orientação sobre requisitos de tarefas
3:1992 4:1998 7:1998 8:1997 9:2000	Considerações ergonômicas relativas ao <i>hardware</i>
5:1998 6:1999	Considerações ergonômicas relativas a estações de trabalho e ambientes
10:1996 11:1998 12:1998 13:1998 14:1997 15:1997 16:1999 17:1998	Considerações ergonômicas relativas ao <i>software</i>

A seção introdutória de quaisquer das partes do padrão ISO 9241 (e.g., ISO 9241-14:1997(E) [ISO97a]) deixa clara a essência de cada uma das partes, resumida como informação complementar no Quadro 26. Tal informação serve de base para a seleção de partes do padrão concernentes a contextos avaliatórios específicos. Tome-se como exemplo a análise de

conformidade do produto-alvo desta pesquisa com o padrão *ISO 9241*. Diferentes níveis de decisão nortearam o processo de inspeção aqui descrito. Como se sabe, o foco central deste trabalho é a confrontação de três enfoques avaliatórios, centrados na usabilidade de produtos interativos usuário-computador – *mensuração do desempenho*, *mensuração da satisfação* e *inspeção de padrões* – com propósitos de investigação da influência do atributo central⁷³ de cada um dos enfoques sobre a usabilidade de interfaces usuário-computador.

O primeiro nível de decisão envolveu a escolha do constituinte computacional do processo interativo: o *software* da aplicação, o *hardware* ou a interface (independentemente do produto-alvo, ainda desconhecido) sobre o qual recairia o foco do processo avaliatório. Em virtude do tema pesquisado, a decisão recaiu sobre a interface do produto com o usuário.

Tendo em vista que a inspeção de conformidade como o padrão *ISO 9241* não era o foco central do trabalho, a interface de qualquer aplicativo de *software* (e.g., um editor de texto, uma planilha eletrônica, um SIG) poderia *a priori* ser o alvo da inspeção. Entretanto, o desenvolvimento da pesquisa no âmbito do *Departamento de Engenharia Elétrica da UFPB* direcionou a escolha do alvo para um produto de *software* de aplicação mais específica na engenharia, que satisfizesse simultaneamente a condição de ser utilizado por um contingente de usuários (principiantes, intermediários e experientes) adequado às atividades de mensuração do desempenho e a sondagem da satisfação contempladas no contexto da pesquisa. A conjunção destas condições implicou a tomada da segunda decisão: a escolha do *MATLAB* como produto de teste.

O terceiro nível de decisão envolvendo a análise de conformidade do produto a um padrão considerou o fato da escolha recair sobre um produto internacional, assim como o fato de não se dispor atualmente de uma norma brasileira nem de um padrão regional concernente a aspectos da interação homem-máquina, conforme reportado na seção 5.3. A escolha de um padrão internacional e mais especificamente do padrão *ISO 9241* não deveu-se apenas à falta de opções de padrões de maior expressividade para processos interativos homem-máquina e, mais especificamente, usuário-computador. Mesmo antes da proposição do tema da pesquisa, o processo de revisão e atualização bibliográficas já havia contemplado o padrão *ISO 9241*, por conta de sua expressividade enquanto padrão internacional. A publicação da *Parte 11* como *IS* em 1998 foi decisiva para a escolha, tendo em vista que até então nenhum outro padrão focalizara a definição de *usabilidade* nem a orientação para sua mensuração.

O último nível de decisão endereçou à seleção das partes do padrão escolhido pertinentes ao contexto da análise de conformidade. Como neste ponto a revisão bibliográfica realizada já propiciara o esclarecimento da essência das diferentes partes do *ISO 9241*, reforçado pela aquisição da *Parte 11*, o fator decisivo para a escolha das demais partes a serem utilizadas foram os modos de interação usuário-produto oferecidos pela interface da versão 5.3.0 do *MATLAB*.

O Anexo D apresenta as principais características do referido produto, dentre as quais figuram os modos de interação - *menus*, *linguagem de comandos* e *manipulação direta*, respectivamente tratados nas *Partes 14, 15 e 16* do padrão *ISO 9241*.

⁷³ *Desempenho* para a mensuração do desempenho, *satisfação subjetiva do usuário* para a mensuração da satisfação e *grau de adequação a um padrão* para o enfoque inspeção de padrões.

A *Parte 11* foi escolhida para dar suporte ao processo avaliatório como um todo, a partir da orientação fornecida sobre a avaliação da usabilidade de interfaces em termos de medidas de desempenho e satisfação do usuário. As demais partes são direcionadas para o processo de inspeção dos modos de interação oferecidos pelo produto avaliado à luz das recomendações da ISO.

5.5.1 Aplicação das Partes 14, 15 e 16 do Padrão ISO 9241

5.5.1.1 Adequação de Diálogos via Menus, Linguagens de Comandos e Manipulação Direta

Da bibliografia consultada sobre projeto e avaliação de interfaces (e.g., ISO [ISO97a, ISO97b, ISO99], Mayhew [Mayh92, Mayh99], Nielsen [Niel93b, Niel00], Salvendy [Salv97]), sabe-se que os diálogos intermediados por interfaces de *software* podem assumir diferentes estilos, com o propósito fundamental de serem particularmente apropriados para uma ou mais condições que refletem diferentes aspectos relativos ao usuário, à tarefa e ao sistema. O Quadro 27 apresenta de forma sucinta condições concernentes a diferentes aspectos do contexto do usuário, da tarefa e do sistema, segundo a visão da ISO [ISO97a, ISO97b, ISO99]. Tais condições contribuem para o direcionamento do projeto de diálogos em interfaces usuário-computador envolvendo os estilos de interação via menus, linguagem de comandos e/ou manipulação direta.

A competitividade do mercado internacional vem estimulando cada vez mais os desenvolvedores de aplicações de *software* de uso genérico e específico a procurarem atender a todas as categorias de usuários, o que tem implicado a geração de produtos com interfaces mais flexíveis em nível do diálogo com o usuário, dotados de diferentes estilos de interação. Os modos de interação observados no Quadro 27 são precisamente aqueles oferecidos pelo produto inspecionado, que busca atender as necessidades de comunicação do contingente de usuários experientes com diálogos via linguagem de comandos, enquanto aos usuários intermediários e principiantes são reservados, respectivamente, os estilos de interação via menus e manipulação direta. Evidentemente, esta é uma leitura simplista da interface do *MATLAB v.5.3*, assim como é simplista afirmar que um dado estilo de interação se destina a uma dada categoria de usuários. Parece mais sensato crer que a diversidade de estilos enriquece a interação entre o usuário e a aplicação que os oferece.

A comunicação entre humanos e sistemas computacionais via terminais de visualização pode ser bastante dinâmica, estimulando a prática, em maior ou menor grau, da criatividade da qual os seres humanos são naturalmente dotados. Por outro lado, o aumento da flexibilidade de comunicação com o usuário através de sua interface pode conferir ao produto maiores chances de aceitação no mercado consumidor e, sem dúvida, confere ao usuário a possibilidade de experimentação dos recursos de diálogo oferecidos e de escolha daqueles que se prestam mais adequadamente à realização de tarefas pertinentes aos mais variados contextos.

Partindo de tais premissas, um mesmo usuário poderá interagir com um sistema através dos diferentes recursos que este lhe ofereça, desde que tal diversificação de uso de estilos lhe seja conveniente e lhe propicie níveis de conforto, segurança e desempenho adequados à realização de suas atividades.

Quadro 27 – Aspectos do contexto do usuário e da organização, das tarefas e do sistema para diferentes estilos de interação, segundo a visão da ISO.

Alvo	Estilo de Interação		
	Manipulação Direta (<i>Parte 16</i>)	Menu (<i>Parte 14</i>)	Linguagem de Comandos (<i>Parte 15</i>)
Usuários e Organização	Mediação da comunicação de indivíduos pouco hábeis em leitura e escrita, mas com habilidades sensorimotoras necessárias à manipulação direta de ícones	Mediação da comunicação de indivíduos pouco hábeis em digitação com a aplicação	Mediação da comunicação de indivíduos com boa habilidade em digitação com a aplicação
	Otimização do desempenho do usuário a partir sugestões visuais que facilitem a memorização	Mediação da interação de indivíduos pouco familiarizados com tecnologias computacionais e com linguagens de comandos	Mediação da interação de indivíduos bastante familiarizados com tecnologias computacionais e linguagens de comandos
	Otimização do desempenho de indivíduos a partir do uso de representações gráficas ao invés de descrições textuais	Facilidade da interação de indivíduos pouco familiarizados ou inexperientes com a aplicação	Mediação da interação de indivíduos bastante familiarizados com a aplicação (uso frequente)
	Minimização de necessidades de treinamento	Minimização de necessidades de treinamento	Possibilidade de treinamento no uso de linguagem de comandos
Tarefas	Existência de metáforas para as aplicações	Inexistência de metáforas para as aplicações	Inexistência de metáforas para as aplicações
	Complexidade dos atributos de objetos associados às tarefas e dificuldade conseqüente de conversão em linguagem usual, porém passível de conversão em linguagem gráfica	Uso esporádico de recursos relativos às tarefas	Arbitrariedade da ordem de entrada de dados e opções, i.e. possibilidade de entrada da informação sem a necessidade de execução de uma seqüência rígida de ações
		Necessidade usual de orientação quanto à variedade de opções oferecidas pelo sistema	
	Dificuldade de pré-determinação das seqüências de ações relativas às tarefas, implicando a necessidade de maior flexibilidade	Relevância da limitação do número de escolhas para a execução das tarefas em contextos específicos	Impossibilidade de previsão das escolhas das ações executadas pelo usuário no contexto do diálogo
	Maior controle do usuário sobre os objetos associados às tarefas	Necessidade de dispositivos de entrada diferentes de teclados (e.g., dispositivos de seleção e apontamento)	Extensibilidade requerida (i.e. criação de novos comandos ou seqüências de comandos)
	Dificuldade de descrição e memorização das entradas, compensada pela facilidade de visualização	Dificuldade de memorização de seqüências muito extensas de comandos necessárias à execução das tarefas, possíveis de serem mapeadas em árvores de opções selecionáveis	Necessidade de seleção ou acesso rápido a funções específicas do sistema
	Facilidade de execução das tarefas a partir de do uso de objetos gráficos visíveis e diretamente manipuláveis		
	Necessidade de transformação de atributos visuais associada às tarefas	Necessidade da visualização constante de opções correntes (<i>default</i>) para a eficácia da execução das tarefas	Necessidade de maior controle do usuário sobre a interação e sobre a navegação através das funcionalidades oferecidas pelo sistema
Necessidade de tratamento das entidades relativas às tarefas como objetos individuais			
Sistema	Resolução de vídeo e dispositivos de entrada de dados adequados para a manipulação direta de objetos de forma detalhada e precisa	Tempo de resposta adequado para a validação de escolhas feitas em seqüências de opções (e.g., em torno de 2 s)	
	Potencialidades técnicas suficientes para a geração eficaz de representações gráficas de objetos	Teclado com recursos limitados	
	Tempos de resposta suficientemente curtos para o retorno imediato às operações de manipulação direta executadas pelo usuário		

É mesmo possível que o usuário venha a empregar, com níveis similares de eficiência, efetividade e satisfação, os diferentes recursos estilísticos do diálogo com a aplicação, assim como os ambidestros fazem uso indistinto das duas mãos para executarem as mesmas tarefas.

Conforme a ISO [ISO97a, ISO97b, ISO99], a aplicabilidade das recomendações apresentadas, respectivamente, nas *Partes 14, 15 e 16* sobre diálogos fundamentados em menus, linguagens de comando e/ou manipulação direta será tanto mais efetiva quanto maior for o número de condições satisfeitas dentre aquelas apresentadas no Quadro 27.

5.5.1.2 Aplicabilidade das Recomendações da ISO

A *Parte 14* do padrão ISO 9241 [ISO97a] oferece recomendações relativas ao projeto ergonômico da interação usuário-computador fundamentada em menus, estruturadas em quatro seções, assim discriminadas: **5 (Estrutura de menus)**, **6 (Navegação em menus)**, **7 (Seleção e execução de opções)** e **8 (Apresentação de menus)**. Nesta parte, as sub-seções da seção **7 (Seleção e execução de opções)** fornecem recomendações concernentes a um método de seleção particular, contendo adicionalmente notas no início da sub-classe relativas à adequação daquele método a um contexto específico.

No tocante à *Parte 15* [ISO97b], as recomendações, concernentes ao projeto ergonômico da interação usuário-computador via linguagem de comandos, são também estruturadas em quatro seções, a saber: **4 (Estrutura e sintaxe)**, **5 (Representação de comandos)**, **6 (Considerações de entrada e saída)** e **7 (Retorno de informações e ajuda)**.

Por sua vez, as seções **5 (Metáforas)**, **6 (Manipulação de objetos)**, **7 (Recomendações adicionais para a manipulação direta de objetos textuais)**, **8 (Recomendações adicionais para a manipulação direta de janelas)** e **9 (Recomendações adicionais para a manipulação direta de ícones de controle)** contidas na *Parte 16* do padrão ISO 9241 [ISO99] oferecem recomendações para o projeto ergonômico de interfaces usuário-computador com diálogos via manipulação direta.

A Fig. 12 ilustra o formato das recomendações individuais contidas nas diferentes partes do padrão ISO 9241, através de um excerto da sub-seção 6.8 da *Parte 15* [ISO97b].

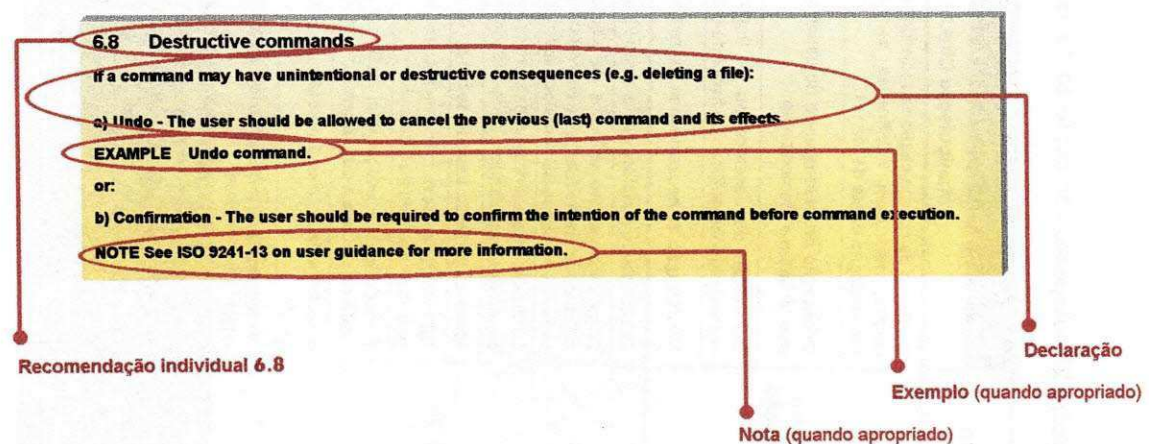


Fig. 12 – Formato das recomendações individuais do padrão ISO 9241.

Mantendo a integridade e consistência em nível da aplicabilidade das recomendações, a ISO [ISO97a, ISO97b, ISO99] orienta que as recomendações individuais necessárias para atingir os objetivos gerais deverão ser aplicadas no contexto específico para o qual são relevantes (e.g., tipos particulares de usuários, tarefas, ambientes e tecnologias). Os exemplos fornecidos para várias das recomendações detalham, em geral, uma implementação que incorpora a recomendação, alguns dos quais também indicam soluções preferenciais.

De acordo com a ISO [ISO97a, ISO97b, ISO99], deve-se avaliar as recomendações individuais por sua aplicabilidade e, caso se julgue serem aplicáveis, implementá-las no diálogo relevante via estilo de interação pertinente, a menos que se constate que fazê-lo implica desvios dos objetivos de projeto ou resulta em uma degradação global da usabilidade da interface e, por extensão, do produto. Sendo constatada a aplicabilidade da recomendação, esta deverá ser avaliada na ordem apresentada na seção ou sub-seção relevante.

Para fins de avaliação de produtos de *software*, ao julgar pertinente a adoção de uma recomendação, a ISO instrui o avaliador a (i) *investigar o produto*, o que implica uma inspeção da conformidade do produto às partes do padrão ISO 9241 consideradas; ou (ii) *observar usuários representativos no contexto de execução de tarefas via sistema de menus*, o que, dito em outras palavras, implica a condução de testes de usabilidade.

5.1.1.3 Síntese dos Aspectos Interativos Abordados pelas Partes 14, 15 e 16 do Padrão ISO 9241

Como já foi anteriormente ressaltado, as *Partes 14, 15 e 16* do padrão ISO 9241 tratam de diversos aspectos da interação usuário-computador relativos, respectivamente, a diálogos via menus, linguagens de comandos e manipulação direta. Esses aspectos aparecem sumariados sob a forma de diagramas de blocos nas páginas seguintes (Figs. 13, 14 e 15), onde também foram discriminados os tópicos abordados no contexto de cada aspecto, com o propósito de facilitar tarefas de consulta às seções da referidas partes do padrão por projetistas e/ou avaliadores de interfaces usuário-computador fundamentadas nos estilos de diálogo mencionados.

Quando se utiliza diretrizes de projeto, guias de estilo, heurísticas ou padrões para auxiliarem atividades de projeto de interfaces, costuma-se confrontar as recomendações oferecidas por estes documentos com as diversas condições concernentes a diferentes aspectos do contexto real (i) do usuário do produto para o qual a interface será desenvolvida; (ii) das tarefas que serão mediadas por tal interface; e (iii) do sistema com o qual a interface servirá de elo de comunicação com o usuário.

Via de regra, as heurísticas costumam ser apresentadas como listas sucintas de princípios de projeto que ocupam, no máximo, três páginas. Por outro lado, os padrões, as coletâneas de diretrizes de projeto e os guias de estilo são documentos que contêm usualmente dezenas de páginas, um inconveniente para projetistas de interfaces, que alegam principalmente o tempo reduzido disponível para a leitura de documentos volumosos e os conflitos que encontram entre as seções de recomendações relativas a diferentes aspectos da interface a ser projetada. Estes

inconvenientes também se reproduzem em atividades de avaliação fundamentadas na análise da conformidade de um projeto a diretrizes de projeto, guias de estilo, heurísticas ou padrões.

A fim de facilitar o trabalho de projetistas na análise da aplicabilidade das recomendações ISO ao contexto de projeto, assim como para auxiliar os avaliadores em atividades de inspeção de conformidade de produtos ao padrão internacional ISO 9241, ao final de cada parte do padrão o primeiro anexo acrescentado pela ISO (denominado *Anexo A*) apresenta procedimentos para análise da aplicabilidade (destinados mais especificamente a projetistas) e/ou da adoção (destinados mais especificamente a avaliadores) das recomendações.

Embora apresente procedimentos destinados a propósitos de projeto e/ou avaliação, a ISO ressalta que a equipe de projeto e/ou avaliação poderá desenvolver outro procedimento que se adeque melhor ao seu contexto de trabalho (ambiente físico, recursos materiais e humanos, etc.).

No caso específico das partes utilizadas nesta pesquisa, a ISO [ISO97a, ISO97b, ISO99] alerta que se um produto anunciar ter adotado recomendações contidas em uma ou mais das *Partes 14, 15 e 16* do padrão *ISO 9241*, este deverá especificar o procedimento empregado no estabelecimento dos requisitos destinados ao desenvolvimento e/ou à avaliação dos menus, linguagem de comandos e manipulação direta, respectivamente.

Os diagramas apresentados no Anexo E foram construídos antes do processo de inspeção conduzido nesta pesquisa, com o propósito de facilitar a localização de recomendações específicas por projetistas e avaliadores.

Organizados conforme as estruturas das *Partes 14, 15 e 16* do padrão *ISO 9241*, os referidos diagramas são compostos por blocos que apresentam, no canto superior esquerdo, um retângulo contendo o número da seção ou sub-seção da parte considerada. No canto inferior direito de cada bloco, um círculo evidencia o número da página da parte considerada na qual se encontra cada seção ou sub-seção.

Além destes recursos de localização, os blocos pertencentes a cada seção e respectivas sub-seções aparecem cromaticamente discriminados dos demais.

5.5.2 Considerações sobre a Aplicabilidade e Adoção das *Partes 14, 15 e 16* do Padrão *ISO 9241*

5.5.2.1 Avaliação de Produtos

A aplicabilidade de um estilo de interação ou de uma combinação de estilos dentre aqueles contemplados pelas *Partes 14, 15 e 16* ao contexto avaliado, nos casos de avaliação de projetos, será tanto mais efetiva quanto mais conformidade o projeto apresentar com uma das partes ou uma combinação das partes mencionadas do padrão *ISO 9241*, i.e. quanto mais recomendações forem satisfeitas. No tocante à avaliação de produtos, a conformidade com as partes referidas do padrão *ISO 9241* será tanto maior

quanto maior for a adoção às recomendações nelas contidas, i.e. quanto maior for o número de recomendações verificadas pelo avaliador.

Segundo o padrão ISO 9241 [ISO97a, ISO97b, ISO99], o procedimento de avaliação deverá fundamentar-se em uma análise de usuários típicos, realizando tarefas típicas e críticas em ambientes típicos. Os procedimentos de avaliação de diálogos via menus, linguagem de comandos e/ou manipulação direta se enquadram usualmente em um dos dois contextos apresentados a seguir:

1. Usuários e tarefas são conhecidos, de modo que os avaliadores inspecionam o produto sob o prisma dos estilos de diálogo oferecidos ou observam usuários representativos do produto durante a realização de tarefas típicas e críticas em ambientes típicos e envolvendo os estilos de interação em questão;

2. Usuários e tarefas específicos não são conhecidos, de modo que os avaliadores inspecionam todos os menus, linguagem de comandos e/ou tipos de manipulação direta do produto considerado.

A ISO também recomenda que os usuários das *Partes 14, 15 e/ou 16* do padrão ISO 9241 demonstrem como o projeto ou o produto de interesse foi inspecionado segundo as recomendações das partes referidas do padrão listando:

(i) *menus, comandos e/ou tipos de manipulação direta avaliados* (e.g., todos os menus, comandos e/ou tipos de manipulação direta ou um sub-conjunto deles associados à tarefa observada);

(ii) *o método adotado para julgar a aplicabilidade das recomendações ao projeto avaliado ou o método adotado para julgar a adoção das recomendações pelo produto avaliado*; e

(iii) *os resultados do procedimento de avaliação*.

5.5.2.2 Aplicabilidade

A aplicabilidade das recomendações contidas em cada uma das partes mencionadas fundamenta-se em dois fatores, a saber:

i. **Se a declaração condicional, caso inclusa como parte da cláusula, for verdadeira**, tendo em vista que uma determinada recomendação só será (ou não) considerada aplicável quando a declaração condicional se for (ou não) verdadeira (e.g., se o tempo de busca não for importante, a recomendação 5.1 da *Parte 14* não será aplicável).

ii. **Se as limitações relativas ao usuário, à tarefa, ao ambiente e à tecnologia o permitirem**, i.e., se for possível considerar as recomendações apesar das limitações do processo de delineamento do perfil da comunidade usuária; da variabilidade das tarefas; do nível de ruído ou de iluminação do ambiente; da resolução do dispositivo de visualização utilizado; ou da

disponibilidade de dispositivos de apontamento.

Portanto, é conveniente considerar que *uma determinada recomendação será aplicável se o ambiente de projeto envolver características do usuário, das tarefas ou da tecnologia utilizada, levadas em conta naquela recomendação, por exemplo:*

(1) *se for permitida a seleção de uma opção de menu via dispositivo de apontamento e seleção, as recomendações condicionais da seção 7.5 (Pointing) deverão ser analisadas, a fim de que se determine se sua aplicabilidade é pertinente ao contexto avaliado;*

(2) *se não houver necessidade de implementar facilidades de rotação dos objetos visualizados na interface avaliada, as recomendações condicionais da seção 6.5 (Rotating) da Parte 16 deverão ser ponderadas, a fim de que se considere ou não sua aplicabilidade ao contexto avaliado.*

5.5.2.3 Métodos para a Determinação da Aplicabilidade e/ou Adoção das Recomendações

A ISO aponta como mais apropriados para a determinação da aplicabilidade de uma dada recomendação os seguintes métodos:

- (i) Análise da Documentação do Sistema (*System Documentation Analysis*);
- (ii) Evidência Documentada (*Documented Evidence*);
- (iii) Observação (*Observation*);
- (iv) Avaliação Analítica (*Analytical Evaluation*); e
- (v) Avaliação Empírica (*Empirical Evaluation*).

Por outro lado, para a determinação da adoção de uma dada recomendação, a ISO considera mais apropriados os seguintes métodos:

- (i) Mensurações (*Measurements*);
- (ii) Observação (*Observation*);
- (iii) Evidência Documentada (*Empirical Evaluation*);
- (iv) Avaliação Analítica (*Analytical Evaluation*); e
- (v) Avaliação Empírica (*Empirical Evaluation*).

Para quaisquer dos propósitos acima mencionados, são descritos detalhada e respectivamente nas seções A.3 e A.5 do Anexo A das partes do padrão ISO 9241 consideradas nesta pesquisa todos os métodos apontados para a determinação da aplicabilidade e/ou da adoção de recomendações particulares.

O Quadro 28 apresenta uma síntese descritiva dos métodos apontados pela ISO para a determinação da aplicabilidade e/ou adoção de recomendações particulares das partes do padrão internacional ISO 9241. As células destacadas em fundo cinza indicam os métodos não apropriados para a execução quer da aplicabilidade, quer da adoção de recomendações específicas de quaisquer das partes anteriormente referidas a um contexto de projeto e/ou de avaliação.

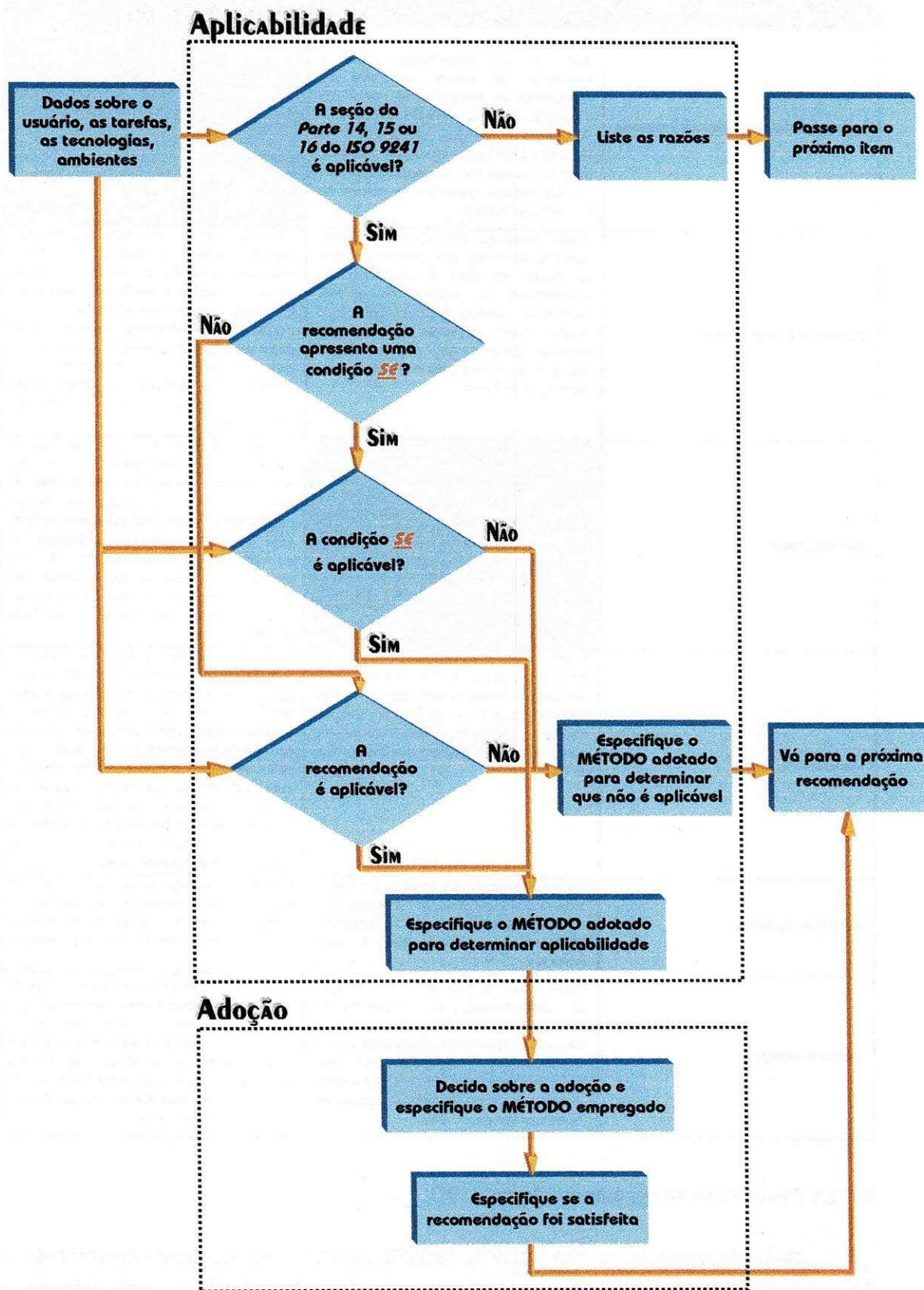
Quadro 28 – Síntese descritiva dos métodos apontados para a determinação da aplicabilidade/adoção de recomendações particulares do ISO 9241.

Método	Aplicabilidade	Adoção
Análise da Documentação do Sistema	Análise de quaisquer documentos que descrevam as propriedades gerais e específicas do sistema em nível de estruturação da hierarquia de menus, da linguagem de comandos e/ou dos tipos de manipulação direta (e.g., especificações de requisitos do sistema e do usuário, manuais, guias do usuário), com o propósito de verificar se determinadas condições se aplicam ao contexto considerado.	
Evidência Documentada	Análise de qualquer informação documentada relevante sobre requisitos ou características da tarefa, do fluxo do trabalho, das competências ou aptidões do usuário, convenções ou referências existentes sobre o usuário, dados de testes do projeto de sistemas similares, etc, que respalde a aplicabilidade de uma dada recomendação ao contexto de interesse.	Consulta qualquer informação documentada relevante relativa à adoção de menus, linguagem de comandos e manipulação direta às recomendações condicionais apropriadas, e.g. convenções ou referências existentes sobre o usuário, dados de testes de protótipos, dados de testes do projeto de sistemas similares, etc. que evidenciem a adequação das recomendações aplicadas às categorias de usuários e tarefas relevantes à aplicação inspecionada.
Mensurações		Determinação a partir de medição, estimação ou cálculo de uma ou mais variáveis concernentes a propriedades da interface inspecionada, do ponto de vista dos estilos de diálogo considerados (menus, linguagem de comandos e/ou manipulação direta), e.g. tempo de resposta [ISO97a, ISO97B, ISO99], área ativa em telas sensíveis ao toque, cálculo do dimensionamento do número de grupos de opções a partir da fórmula apresentada [ISO97a].
Observação	Exame ou inspeção da estrutura de menus, da estruturação da linguagem de comandos e/ou dos recursos de manipulação direta oferecidos pelo sistema, a fim de estabelecer relações entre determinadas propriedades observáveis e recomendações contidas nos documentos da ISO.	Exame ou inspeção da estrutura de menus, da estruturação da linguagem de comandos e/ou dos recursos de manipulação direta oferecidos pelo sistema, a fim de confirmar a conformidade de determinadas condições observáveis com as recomendações correspondentes da ISO (e.g., títulos de menus consistentemente alinhados pela esquerda ou centralizados, nomes de comandos não excedem sete caracteres, progresso do dimensionamento de objetos via manipulação direta ilustrados por uma seqüência de indicações visuais).
Avaliação Analítica	Julgamentos abalizados sobre as propriedades dos estilos de diálogo via menus, linguagem de comandos e/ou manipulação direta, usualmente feitos por especialistas em situações nas quais o sistema só existe em termos de documentos de projeto, quando não há populações usuárias disponíveis ou quando restrições em termos de recursos e/ou tempo são impostas ao processo de avaliação.	
Avaliação Empírica	Procedimentos de teste para a determinação da aplicabilidade de recomendações envolvendo usuários finais representativos. Tal estratégia é usualmente considerada quando há a disponibilidade de protótipos do sistema inspecionado ou do próprio sistema e quando também há a disponibilidade de populações usuárias potenciais ou reais.	Procedimentos de teste para a determinação da adoção de recomendações condicionais que envolvem usuários finais representativos. Tal estratégia é usualmente adotada quando há a disponibilidade de protótipos do sistema inspecionado ou do próprio sistema e quando também há a disponibilidade de populações usuárias potenciais ou reais.

5.5.2.4 Procedimento recomendado pela ISO

Segundo sugestão da ISO [ISO97a, ISO97b, ISO99], o procedimento apresentado no fluxograma da Fig. 13 pode ser empregado em processos de inspeção da aplicabilidade de recomendações contidas nas *Partes 14, 15 e 16* do padrão ISO 9241 a interfaces projetadas com

base em estilos de diálogo via menus, linguagem de comandos e/ou manipulação direta e/ou em processos de inspeção da adoção de tais recomendações por interfaces já implementadas.



No tocante à aplicabilidade de recomendações condicionais, cada recomendação condicional apresenta uma condição se, seja na própria declaração do item (e.g., *item 6.1.1*) ou implícita no título de uma sub-seção (e.g., *sub-seção 8.4*), de modo que a aplicabilidade da declaração se deverá ser determinada a partir dos métodos propostos para o teste de veracidade da condição se (e.g., no *item 5.1.4* da *Parte 14*, o método da *evidência documentada*, da *avaliação analítica* ou da *avaliação empírica* se afiguram apropriados para determinar se um tempo de busca rápido é importante no contexto considerado).

Além disso, quando houver um conjunto de recomendações condicionais opcionais (e.g., *5.1.1*, *5.1.2* e *5.1.3* ou *5.2.1* e *5.2.2* da *Parte 14*, a escolha da abordagem aplicável também deverá ser determinada a partir dos métodos propostos.

Quanto à adoção de recomendações condicionais, para cada recomendação aplicável, a adoção da referida recomendação deverá ser determinada a partir dos métodos propostos (e.g., se o *item 5.1.4* da *Parte 14* é aplicável, dever-se-á adotar o método da *avaliação analítica* ou *avaliação empírica* para determinar que tantos níveis e opções quantos forem possíveis serão integrados em um único painel de menu).

Em geral, as recomendações não condicionais do tipo "*se declaração*" são apropriadas para qualquer sistema de menus, linguagem de comandos ou manipulação direta, embora alguns dos sub-itens de tais recomendações só sejam aplicáveis a sistemas que incorporem os recursos mencionados na declaração, e.g. se o sistema considerado empregar teclas de funções para seleção das opções, as recomendações não condicionais relativas a esta modalidade de interação serão aplicáveis.

No que diz respeito à adoção de recomendações não condicionais do tipo "*se declaração*", é necessário informar sobre a adoção à recomendação considerada, conforme descrito para recomendações condicionais. Para a determinação da adoção à recomendação "*oferecer um mecanismo simples de retorno ao menu inicial*", no contexto da inspeção de um determinado sistema, são apropriados tanto o método da *avaliação analítica* quanto o da *avaliação empírica*. Caso haja razões válidas para a não adoção da recomendação proposta, tanto as razões quanto a solução de projeto adotada deverão ser registradas do relatório da inspeção.

5.5.3 Uso das Listas de Inspeção (*Checklists*) das *Partes 14, 15 e 16*

Ao final do *Anexo A* de cada uma das partes do padrão *ISO 9241*, há uma lista de inspeção (*checklist*) estruturada em formato tabular (*Table A.1*). Conforme mencionado em sub-seção anterior, esta representação tabular sumaria as recomendações contidas na parte considerada do padrão, além de apresentar campos que auxiliam (i) projetistas, na inspeção da aplicabilidade das recomendações condicionais daquela parte do *ISO 9241* a projetos específicos; e (ii) avaliadores de sistemas resultantes da implementação de projetos, na verificação da adoção das recomendações pelos projetistas. Muitas das recomendações condicionais permitem diversas soluções alternativas, motivo pelo qual a lista de inspeção emprega conectores *AND (E)* e *OR (OU)*, que são apresentados, em alguns casos, como *AND/OR (E/OU)*, quando as escolhas não são mutuamente exclusivas.

	Recomendações		Aplicabilidade						Adoção						Comentários (incluindo fontes)	
			Resultado		Método usado				Resultado		Método usado					
	S	N	A	E	O	AA	AE	MD	W	O	D	AA	AE	MD	P	F
5 Informação Geral																
5.1 Metáforas																
5.1.1 Disponibilização de um contexto estrutural																

Fig. 14 – Excerto da lista de inspeção contida no Anexo A das Partes 14, 15 e 16 do ISO 9241.

Para a realização deste trabalho, as listas de inspeção das Partes 14, 15 e 16 do ISO 9241 foram traduzidas e adaptadas para a língua portuguesa, sendo reproduzidas no Anexo F, ao final deste documento. Conforme se observa na Fig. 14, um excerto das seis primeiras linhas da representação tabular contida no Anexo F3, a lista de inspeção existente ao final de cada uma das partes utilizadas do padrão ISO 9241 contém 18 colunas, assim distribuídas:

i. **Coluna de Recomendações** (*Recommendations column*)

Apresenta as recomendações condicionais da parte considerada do ISO 9241 em uma versão sumariada, concatenadas através de conectores lógicos, separadas por sub-itens e numeradas conforme aparecem no texto das recomendações. Tal estratégia agiliza atividades de consulta nos casos em que o projetista/avaliador não dispõe de tempo para uma leitura mais prolongada do texto integral. A fim de evitar indecisão ou o preenchimento incorreto da tabela, as células que não devem ser assinaladas foram pintadas em cinza escuro.

ii. **Colunas para Registro da Aplicabilidade** (*Recommendations column*)

As duas primeiras colunas relativas à aplicabilidade devem ser empregadas para o registro dos resultados da determinação da aplicabilidade mediante a aposição de um sinal de verificação (✓) na coluna **S** (*Sim*) ou **N** (*Não*). Além destas duas colunas, o bloco referente à aplicabilidade também indica quais dos métodos de avaliação da aplicabilidade foi adotado, oferecendo células em branco para que o projetista/avaliador assinale para cada uma das recomendações condicionais, o método adotado.

As células associadas aos métodos que não são relevantes para determinadas recomendações aparecem preenchidas em cinza escuro. As abreviações originais empregadas nas colunas referentes aos diferentes métodos recomendados pela ISO foram codificadas em português como:

- **A** ⇨ Análise da Documentação do Sistema
- **E** ⇨ Evidência Documentada
- **O** ⇨ Observação
- **AA** ⇨ Avaliação Analítica
- **AE** ⇨ Avaliação Empírica

Caso um método diferente seja adotado, a célula correspondente na coluna **MD** deverá ser assinalada e o método descrito na última coluna (**Comentários**).

iii. Colunas para Registro da Adoção (*Adherence column*)

As duas últimas colunas deste bloco devem ser empregadas para registro do resultado da determinação da adoção mediante a aposição de um sinal de verificação (✓) quer na coluna **P** (*Passou*) ou **F** (*Falhou*).

Por outro lado, as primeiras colunas deste bloco indicam quais dos métodos de aplicabilidade são relevantes para cada uma das recomendações condicionais, oferecendo espaços para que o projetista/avaliador assinale o método adotado, similarmente às colunas correspondentes do bloco da aplicabilidade. Aqui também as células associadas aos métodos não relevantes para determinadas recomendações aparecem preenchidas em cinza escuro. As colunas referentes aos diferentes métodos recomendados pela ISO são rotuladas com o seguinte código:

- ◉ **M** ⇨ Mensuração
- ◉ **ED** ⇨ Evidência Documentada
- ◉ **O** ⇨ Observação
- ◉ **AA** ⇨ Avaliação Analítica
- ◉ **AE** ⇨ Avaliação Empírica
- ◉ **MD** ⇨ Método Diferente

Similarmente ao que ocorre com o bloco da aplicabilidade, caso um método diferente tenha sido considerado na determinação da adoção de uma dada recomendação, a célula correspondente na coluna **MD** deverá ser assinalada e o método descrito na última coluna (**Comentários**).

iv. Coluna de Comentários (*Comments column*)

Contém células para informações e comentários adicionais relativos à cada uma das recomendações condicionais, sendo também empregadas para registro da fonte da inspeção (e.g., nome do avaliador, título da evidência documentada) e para descrição de "Métodos Diferentes", quando se mostrar pertinente.

Vale a pena ressaltar que as colunas destinadas ao registro da aplicabilidade e adoção das listas de inspeção apresentadas no Anexo F foram preenchidas com os resultados da inspeção de conformidade do *MATLAB v. 5.3.0* às *Partes 14, 15 e 16* do padrão *ISO 9241*. A título de exemplificação, a coluna de comentários a lista de inspeção referente à *Parte 15* do padrão *ISO 9241* aparece preenchida no Anexo F2.

5.5.4 Sumário dos Resultados de uma Inspeção

Os resultados da avaliação devem ser sumariados a partir da computação de um **Percentual de Adoção (PA)**, que é a relação entre o número de sinais de verificação na coluna **S** e o número de sinais de verificação na coluna **P** multiplicado por 100%, representando o percentual de recomendações aplicáveis satisfatoriamente adotadas.

A ISO considera altamente recomendável o relato de todos os dados, i.e., o número de **P** e o número de **S** conjuntamente aos **PA**. Dependendo da complexidade do sistema de menus, a ISO recomenda o preenchimento de uma lista de inspeção para cada menu do sistema e a ponderação dos **PA** relativos a cada menu, para a determinação do **PA** médio do sistema de menus. Entretanto, a ISO também alerta que tal taxa não passa de um indicador numérico, não podendo, pois, ser utilizada como uma medida confiável do grau de adoção de recomendações aplicáveis sem que leve em conta os respectivos pesos dos itens, individualmente e no contexto de uso.

5.6 Comentários Finais

Este capítulo teve o propósito de apresentar e discutir os vários aspectos relativos à metodologia utilizada na inspeção de conformidade do **MATLAB v. 5.3** (*The MathWorks Inc.*) *Partes 11, 14, 15 e 16* do padrão internacional *ISO 9241*. Foram discutidas questões referentes à padronização dos processos interativos usuário-computador, com ênfase na importância e benefícios advindos do uso de padrões em atividades de projeto e na avaliação de interfaces.

Adicionalmente, foram listados e comentados alguns prós e contras das iniciativas de padronização da interação homem-máquina, após o que foi oferecida uma visão sinóptica dos padrões concernentes a diferentes aspectos dos processos interativos homem-máquina, atualmente disponíveis e em desenvolvimento, tanto em nível internacional quanto nacional. Todavia, a ênfase do panorama apresentado recaiu sobre o padrão internacional *ISO 9241*, apresentado mais detalhadamente, em especial as *Partes 11, 14, 15 e 16*, utilizadas nesta pesquisa.

Por fim, foram detalhadas as etapas do procedimento metodológico adotado no processo de inspeção da conformidade do produto selecionado às *Partes 11, 14, 15 e 16* do padrão internacional *ISO 9241*. Vale a pena salientar que os procedimentos descritos na seção anterior representam apenas uma parte da metodologia adotada no enfoque híbrido ou adaptativo considerado neste trabalho. Como se sabe, o referido enfoque envolveu, além da inspeção de conformidade do produto selecionado às *Partes 11, 14, 15 e 16* do padrão internacional *ISO 9241*, a mensuração do desempenho do usuário durante o uso do produto e a sondagem da satisfação do usuário no tocante ao produto.

Similar, em estrutura, a este capítulo, o próximo capítulo descreverá os procedimentos metodológicos referentes aos dois últimos enfoques mencionados no parágrafo anterior.

6

SONDAGEM DA SATISFAÇÃO SUBJETIVA E MENSURAÇÃO DO DESEMPENHO DO USUÁRIO

"The real question before us lies here: do these instruments further life and enhance its values, or not?"

[Lewis Mumford – *Technics and Civilization* (1934)]

6 Sondagem da Satisfação Subjetiva e Mensuração do Desempenho do Usuário	173
6.1 Importância e Benefícios da Sondagem da Satisfação Subjetiva do Usuário ...	175
6.2 Importância e Benefícios da Mensuração do Desempenho do Usuário	176
6.3 Sondagem da Satisfação Subjetiva do Usuário.....	178
6.3.1 Questionários Disponíveis no Mercado para a Sondagem da Satisfação Subjetiva do Usuário de Sistemas Computacionais.....	178
6.3.1.1 QUIS (<i>Questionnaire for User Interaction Satisfaction</i>).....	179
6.3.1.2 ASQ (<i>After-Scenario Questionnaire</i>), PSSUQ (<i>Post-Study System Usability Questionnaire</i>) e CSUQ (<i>Computer System Usability Questionnaire</i>).....	180
6.3.1.3 SUS (<i>System Usability Scale</i>).....	182
6.3.1.4 IsoMetrics.....	184
6.3.1.5 SUMI (<i>Software Usability Measurement Inventory</i>).....	186
6.3.2 Instrumentos de Sondagem Desenvolvidos no Âmbito desta Pesquisa.....	188
6.3.2.1 DePerUSI (<i>Definimento do Perfil do Usuário de Sistemas Interativos</i>)	192
6.3.2.2 OpUS (<i>Opinião do Usuário de Software</i>)	195
6.4 Metodologia Adotada.....	197
6.4.1 Planejamento do Ensaio Avaliatório.....	198
6.4.1.1 Caracterização do Alvo de Estudo.....	198
6.4.1.2 Definição das Metas e Interesses.....	199
6.4.1.3 Caracterização do Universo Amostral.....	199
6.4.1.4 Levantamento dos Usuários de Teste Potenciais.....	199
6.4.1.5 Definição do Modo de Recrutamento dos Participantes.....	199
6.4.1.6 Decisão do Número de Participantes do Ensaio Avaliatório.....	200
6.4.1.7 Seleção das Técnicas de Avaliação.....	200
6.4.1.8 Definição de Indicadores Objetivos e Subjetivos.....	201
6.4.2 Elaboração do Material do Ensaio.....	204
6.4.2.1 Planejamento e Estruturação de Tarefas de Teste.....	204
6.4.2.3 Elaboração da Ficha Cadastral do Participante e do Documento de Aceitação das Condições de Teste.....	208
6.4.2.4 Elaboração do Material Necessário à Condução do Processo Avaliatório.....	208
6.4.2.5 Discussão das Abordagens de Avaliação da Usabilidade do Produto.....	208
6.4.2.6 Validação do Material Elaborado.....	208
6.4.3 Condução do Ensaio e Coleta de Dados.....	209
6.4.4 Tabulação e Análise dos Dados	209
6.4.4.1 Triagem Preliminar dos Dados Coletados.....	209
6.4.4.2 Triangulação dos Dados Coletados.....	210
6.4.4.3 Tabulação e Síntese dos Dados Coletados	210
6.4.4.4 Organização dos Problemas Listados.....	210
6.4.5 Apresentação dos Resultados.....	210

Este capítulo integra os dois enfoques avaliatórios que envolveram a participação de usuários do produto: a *sondagem da satisfação subjetiva* e a *mensuração do desempenho do usuário*. Estes enfoques, juntamente com a inspeção de conformidade do produto-alvo com as *Partes 11, 14, 15 e 16* do padrão internacional ISO 9241 (tratada no capítulo anterior), compõem a tríade de enfoques sobre os quais se fundamenta este trabalho de pesquisa. A razão para a apresentação conjunta da *sondagem da satisfação subjetiva* e da *mensuração do desempenho do usuário* deve-se à sua integração no procedimento metodológico que fundamentou a parte experimental da avaliação da usabilidade do produto-alvo (o *MATLAB v. 5.3 da The MathWorks Inc.*) envolvendo usuários de teste.

Na seção 6.1 (**Importância e Benefícios da Sondagem da Satisfação Subjetiva do Usuário**), retoma-se a abordagem adotada nas seções 2.1 e 3.2.1.2, destacando a relevância da sondagem da satisfação subjetiva do usuário para atividades de projeto e avaliação de interfaces e enfatizando a importância e os benefícios advindos do uso de questionários como instrumentos destinados ao delineamento do perfil e à sondagem da satisfação subjetiva de sistemas computacionais.

A seção 6.2 - **Importância e Benefícios da Mensuração do Desempenho do Usuário** - segue a mesma linha de abordagem da seção anterior, desta feita ressaltando a importância e benefícios advindos da mensuração do desempenho de usuários de sistemas computacionais para atividades de projeto e avaliação de interfaces.

A seção seguinte (**Instrumentos Utilizados na Sondagem da Satisfação Subjetiva do Usuário**), fragmentada em três sub-seções, esboça sucintamente o panorama dos questionários psicométricos disponibilizados atualmente no mercado e descreve os dois instrumentos desenvolvidos e adotados nesta pesquisa para o delineamento do perfil e a sondagem da opinião do usuário de *software*.

A seção 6.4 - **Metodologia Adotada** - descreve detalhadamente os procedimentos de delineamento do perfil e da sondagem da satisfação subjetiva do usuário, aos quais foi intercalada a mensuração do desempenho, compondo a parte experimental da avaliação da usabilidade do produto-alvo que envolveu a participação de usuários de teste.

6.1 Importância e Benefícios da Sondagem da Satisfação Subjetiva do Usuário

Produtos de *software* que agradam ao usuário, deixando-os subjetivamente satisfeitos durante e após as sessões de uso, são mais facilmente aceitos do que aqueles que o irritam, o frustram ou lhe desagradam [Cons99]. Sem dúvida, estes últimos só serão usados em circunstâncias nas quais não há opções, geralmente em nível de disponibilidade no mercado ou possibilidade de aquisição. O desenvolvimento de produtos fundamentado na satisfação subjetiva do usuário transcende os *slogans* publicitários, principalmente quando se trata de aplicações de uso contínuo, visto que os processos interativos usuário-computador que aborrecem, cansam ou desgostam o usuário tendem a reduzir sua eficiência, induzindo-o a concentrar-se menos no trabalho e, por conseguinte, cometer mais erros.

Conforme a discussão conduzida na seção 3.2.1.2, o uso de questionários como instrumentos de delineamento do perfil e sondagem da satisfação subjetiva de usuários de sistemas computacionais tem sido talvez, nas duas últimas décadas, a estratégia mais comum de análise de atitudes, opiniões e preferências empregada por avaliadores de interfaces. Validada a partir dos inúmeros levantamentos usualmente conduzidos por investigadores atuantes do domínio das ciências sócio-humanas, a técnica da sondagem a partir de questionários foi adaptada para o contexto da avaliação de sistemas interativos usuário-computador por apresentar vários atrativos, dentre os quais três se afiguram mais relevantes.

Em primeiro lugar, questionários são instrumentos de coleta de dados que permitem ao investigador, a partir da arguição de indivíduos, a focalização direta do tema de interesse, possibilitando o controle do grau de especificidade almejado para a investigação e a aquisição de dados homogêneos fornecidos, pelo menos em princípio, pela categoria de indivíduos que se insere exatamente no contexto da pesquisa. Deste modo, questionários de usabilidade possibilitam ao avaliador o retorno de informações relativas a atitudes, opiniões e preferências do usuário de sistemas computacionais. Em suma, uma síntese do ponto de vista do usuário, que será tanto mais fidedigna quanto mais bem elaborada for a arguição do usuário sobre o tema de interesse e quanto mais representativa da população-alvo for a amostra selecionada pelo avaliador.

Em segundo lugar, os escores produzidos pela investigação, independentemente dos sistemas, usuários e tarefas considerados, são passíveis de descrição e análise estatística, mediante técnicas de correlação e regressão. Tal atrativo possibilita, portanto, a comparação da usabilidade percebida de sistemas diferentes, a facilidade de uso de um sistema por diferentes categorias de usuários e/ou a facilidade de realização de tarefas diferentes a partir dos recursos oferecidos por um determinado sistema.

Um terceiro atrativo é que o conjunto de dados resultante da sondagem via questionários pode servir de respaldo para inferências estatísticas, a partir da relação entre a amostra inspecionada e a população correspondente.

Além disso, questionários são, em geral, ferramentas de sondagem de aplicação rápida, o que reduz, por conseguinte, os custos envolvidos com a administração e computação dos resultados e possibilita, ao mesmo tempo, a coleta de grandes quantidades de dados.

Adicionalmente, tais dados podem ser usados como base para comparações ou para a demonstração de quais metas quantitativas de usabilidade foram atingidas.

Finalmente, o anonimato do respondente de um questionário (*offline* ou *online*), aliado à ausência de um mediador, durante o preenchimento dos itens, pode implicar a redução ou eliminação dos efeitos de “polarização” das respostas. Por um lado, porque o respondente não se sente compelido a responder aquilo que ele imagina que o mediador gostaria de ouvir ou de ler. Por outro lado, porque não há mediador para “guiar” o processo de formulação de respostas.

Em contrapartida, o uso de questionários apresenta algumas desvantagens. Provavelmente a maior delas seja a reduzida proporção de instrumentos de sondagem devidamente preenchidos e devolvidos. Na realidade, a desvantagem não é *per se* relativa ao baixo número de questionários aproveitáveis devolvidos. Se assim o fosse, o avaliador poderia distribuir uma quantidade de exemplares que compensasse o baixo índice de retorno. O cerne da desvantagem é que o universo de indivíduos que despenderam tempo e esforço para preencher adequadamente e devolver os questionários pode não ser representativo da população de interesse para a pesquisa.

A segunda desvantagem é relativa à subjetividade dos questionários, pois apesar de fornecerem informações sobre atitudes e reações dos respondentes, estes o fazem segundo suas percepções individuais. Por esta razão, questões que dizem respeito ao desempenho do usuário não são usualmente respondidas de modo confiável, não podendo ser empregadas como base para inferências sobre o tema investigado.

Outra desvantagem de menor dimensão é que os questionários são normalmente elaborados de modo a se adaptarem a diferentes contextos de investigação, o que implica a indefinição do sucesso de sua aplicação a contextos nos quais ainda não foram convenientemente validados. Entretanto, questionários bem elaborados podem ser aplicados a gamas mais abrangentes de contextos, com graus de sucesso satisfatórios.

A quarta e última desvantagem aqui mencionada é a necessidade de complementação das informações obtidas através de questionários, já que o ponto de vista do usuário constitui apenas uma das facetas do problema investigado. Quando a meta da investigação é a análise da usabilidade global de um dos componente do *software*, e.g. sua interface, então os dados subjetivos coletados devem ser complementados com dados relativos ao desempenho, efetividade e esforço mental do usuário, o que implica a adoção de outras técnicas (e.g., observação, com ou sem registro de áudio/vídeo).

6.2 Importância e Benefícios da Mensuração do Desempenho do Usuário

Embora as metas qualitativas da usabilidade sejam inegavelmente importantes, conforme visto na seção anterior, não é simples determinar se foram atingidas ao final de um projeto, uma vez que não são usualmente quantificáveis. Em virtude de serem, em geral, especificadas de um modo que dá margem a uma extensa gama de conjecturas, as metas qualitativas não podem ser usadas diretamente como critérios de aceitação durante um ensaio de usabilidade [Mayh99].

Dito de outro modo, não é suficiente que os produtos de *software* agradem ao usuário, deixando-os subjetivamente satisfeitos durante e após as sessões de uso [McGr97]. Embora decisivo para a aceitação de produtos por consumidores, este fato *per si não* assegura a usabilidade dos produtos nem serve de ponto de referência para a comparação objetiva de diferentes produtos desenvolvidos para o mesmo fim, conforme argumentado por alguns pesquisadores da usabilidade de produtos de *software* [Hayh90, Keys90, Niel94b, Kiss95].

Dai a importância das *metas quantitativas de usabilidade*, cuja objetividade e mensurabilidade possibilitam atingir níveis de especificidade necessários à validação de análises fundamentadas em ensaios de usabilidade. Conforme apresentadas no Capítulo 3, as metas quantitativas de usabilidade podem focalizar (i) a *facilidade de uso*, usualmente definida em termos da rapidez potencial, da eficiência e da flexibilidade oferecida por uma interface a um usuário experiente [Cons99, Mayh99]; ou, por outro lado, (ii) a *facilidade de aprendizado*, definida essencialmente em termos da extensão e declividade das curvas de aprendizado de usuários que ainda não atingiram níveis de perícia comparáveis àqueles exibidos por usuários experientes [Mayh99, Pica00]. Enquanto as primeiras estão associadas ao uso do produto avaliado por usuários *experientes*, as últimas envolvem o uso do produto por usuários *principiantes* e *intermediários*. Em ensaios de usabilidade envolvendo participantes com diferentes graus de perícia no uso do produto, é necessária a definição de metas que contemplem tanto a *facilidade de uso* quanto a *facilidade de aprendizado* do produto considerado.

Mayhew [Mayh99] comenta que todas as metas quantitativas de usabilidade ainda podem ser formuladas como (i) *metas absolutas*, quando podem ser traduzidas por indicadores de quantificação absoluta, e.g. o tempo de execução de uma tarefa em minutos ou segundos, o número de ações incorretas cometidas durante a execução de uma tarefa; ou (ii) *relativas*, quando envolvem a comparação de indicadores quantitativos associados a ações do usuário suportadas por produtos de mesma natureza, e.g. tempos de execução de uma tarefa utilizando três produtos similares e concorrentes no mercado de consumo.

Diante do que foi exposto, constata-se que o cerne da questão reside na necessidade de confrontação do usuário com a interface do produto. Indo um pouco além da mera confrontação, é imprescindível avaliar o desempenho do usuário durante o uso do produto. Mais uma vez, faz-se necessário lançar mão dos argumentos de Hix e Hartson [Hix93] e Hackos e Redish [Hack98] para lembrar que o produto é, do ponto de vista do usuário, a interface através da qual este acessa os recursos por ele oferecidos. Assim, não se pode prescindir de uma análise da efetividade e da eficiência com que o usuário interage com o produto, a partir dos mecanismos de comunicação que compõem a interface deste produto, i.e. é indispensável mensurar a efetividade e a eficiência com que a interface do produto acessa os recursos solicitados pelo usuário.

É através dos indicadores de desempenho (vide seção 3.2.1) que a usabilidade de interfaces vem sendo mensurada nos últimos anos [Niel97a], sendo várias as razões que respaldam a mensuração do desempenho de usuários interagindo com produtos de *software*: (i) é necessária para a verificação de produtos, no tocante à adoção de especificações de usabilidade; (ii) dá suporte a tomadas de decisões de projeto, auxiliando nos julgamentos dessas decisões (após sua execução); (iii) possibilita a quantificação das componentes da usabilidade *eficiência* (*efficiency*) e *efetividade* (*effectiveness*) [Jord98, Mayh99, Rask00].

6.3 Sondagem da Satisfação Subjetiva do Usuário

Testes psicológicos constituem, em essência, estratégias padronizadas e objetivas de mensuração de uma amostra do comportamento humano [Anas68, Kira94] e têm sido largamente realizados em pesquisas nas áreas de ciências sociais e da saúde para a mensuração das tendências comportamentais dos indivíduos. Não poderia ser diferente em procedimentos de avaliação de sistemas computacionais, onde a mensuração das tendências comportamentais de usuários finais com relação a um sistema utilizado pode enriquecer o quadro avaliatório do sistema com informações que podem complementar aquelas coletadas a partir da mensuração do desempenho e/ou fornecidas por processos de inspeção do produto por especialistas.

Dolotta *et al.* [Doll76] parecem ter cunhado em seu livro o termo *qualidade percebida pelo usuário* (*user-perceived quality*), largamente utilizado nos dias atuais, enquanto Dzida *et al.* [Dzid78] desenvolveram o primeiro questionário de larga escala especificamente destinado à determinação de uma medida rigorosa da *qualidade percebida pelo usuário*, um conceito que encararam como multidimensional, onde cada dimensão foi associada a um conjunto de propriedades do sistema. A estrutura de sete fatores concebida por Dzida *et al.* [Dzid78] a partir de uma amostra de 100 requisitos de sistemas para a qualidade percebida pelo usuário é um dos precursores da Parte 10 do padrão internacional ISO 9241, *Dialogue Principles*.

Por outro lado, conforme descrito na seção 3.2.1.2, Bailey e Pearson [Bail83] descreveram seu instrumento de mensuração e análise da satisfação de usuários de sistemas computacionais, um questionário que, apesar de só haver sido divulgado em 1983, é praticamente contemporâneo do questionário desenvolvido por Dzida *et al.* [Dzid78], já que foi utilizado em um estudo publicado por Deese [Dees79].

Desde então, conforme a revisão apresentada nas seções 2.1 e 3.2.1.2 deste documento, diversos pesquisadores (e.g., Ives *et al.* [Ives83], Doll e Torkzadeh [Doll88], Chin *et al.* [Chin88], Shneiderman [Shne87], Kirakowski [Kira87, Kira93, Kira96], Lewis [Lewi91a], Brooke [Broo96], Gediga *et al.* [Gedi99]) tem documentado iniciativas de desenvolvimento de instrumentos de sondagem da satisfação subjetiva de usuários de sistemas computacionais.

A seção seguinte apresentará em linhas gerais cinco dos questionários de sondagem da satisfação mais difundidos atualmente no mercado internacional.

6.3.1 Questionários Disponíveis no Mercado para a Sondagem da Satisfação Subjetiva do Usuário de Sistemas Computacionais

Vários autores têm focalizado a atenção quer sobre (i) a elaboração de instrumentos de levantamento e sondagem de opiniões e preferências (e.g., Aiken [Aike97]), (ii) a construção de escalas numéricas e/ou semânticas e a computação da pontuação de listas de inspeção (*checklists*) (e.g., Aiken [Aike96]) ou (iii) o uso de métricas adequadas para a mensuração da satisfação (e.g., Brown [Brow96]).

Paralelamente, diversos instrumentos destinados tanto ao delineamento do perfil de usuários de sistemas computacionais quanto à sondagem de sua satisfação no tocante ao uso de

produtos de *software* tem sido desenvolvido nas duas últimas décadas. As próximas sub-seções sumariam cinco dos instrumentos de sondagem da satisfação subjetiva de usuários de sistemas computacionais que se destacaram em nível mundial na última década.

6.3.1.1 QUIS (*Questionnaire for User Interaction Satisfaction*)

O *QUIS* é um instrumento de sondagem com 80 itens [Harp93] desenvolvido por uma equipe multidisciplinar de pesquisadores do *Human-Computer Interaction Lab (HCIL)* da *University of Maryland*. Segundo Chin *et al.* [Chin88], o *QUIS* foi elaborado para sondar a satisfação, aspecto da usabilidade de produtos de *software*, de um modo padronizado, confiável e válido. A versão preliminar do *QUIS* (v. 2.0), em formato impresso longo ou curto, continha, respectivamente, 90 ou 20 itens associados a uma escala de Likert⁷⁴ de 9 pontos [Shne87, Chin88].

Conforme Slaughter *et al.* [Slau94], desde sua elaboração no formato lápis e papel, têm sido desenvolvidas várias versões eletrônicas deste questionário com a mesma confiabilidade da versão inicial [Harp94]. Embora tais versões tenham contribuído significativamente para o processo de coleta de dados e apresentado vantagens no tocante à configuração dos itens, Harper *et al.* [Harp97] relataram que as versões eletrônicas infelizmente têm provado ser difíceis de manter, distribuir e customizar.

Harper *et al.* [Harp97] relataram que uma atualização do *QUIS* proporcionou a oportunidade de migrá-lo para um formato que proporcionasse sua difusão via Web e, ao mesmo tempo, minimizasse os problemas das versões anteriores. Assim, o *QUIS* v. 7.0, a versão atualizada e expandida da versão 5.5, anteriormente validada, aparece em um formato hierárquico contendo: (i) um questionário demográfico; (ii) seis escalas destinadas à mensuração da reação global do usuário ao sistema; (iii) quatro medidas de fatores específicos à interface (tela, terminologia e retorno de informações, aprendizado e facilidades oferecidas pelo sistema); e (iv) seções opcionais destinadas à avaliação de componentes específicos do sistema (ajuda *online* e impressa, tutoriais *online*, multimídia, acesso à Internet e instalação do *software*).

Cada fator específico à interface e seções opcionais possui uma questão principal, sendo desmembrada em sub-componentes relacionadas, enquanto cada item do *QUIS* é associado a uma escala numérica de 9 pontos delimitada por adjetivos positivos à direita e negativos à esquerda. Adicionalmente, a cada item é oferecida a opção *not applicable* (não aplicável, assim como um espaço que permite a inclusão de comentários).

O *QUIS* 7.0 é atualmente comercializado pelo *University of Maryland Office of Technology Commercialization* nas versões eletrônica e site. O pacote inclui (i) um documento eletrônico contendo todas as seções do questionário, editável conforme necessidades específicas do consumidor; (ii) uma versão HTML do questionário, passível de uso em diversas plataformas, incluindo Mac, Win3.1 e Win95; e (iii) uma coletânea de artigos relevantes detalhando a validação do *QUIS* e algumas de suas aplicações.

⁷⁴ Escala de Likert é uma escala de atitudes, consistindo de um conjunto ímpar de termos de semântica diferencial (e.g., *concordo totalmente, nem concordo nem discordo, discordo totalmente*) através da qual os respondentes indicam seu grau de concordância ou discordância de uma proposição concorrente a algum objeto, indivíduo ou situação [Aike96, Aike97].

São três as modalidades de aquisição: (i) *Commercial license*, U\$ 750 a versão eletrônica e U\$ 1000 a versão Web; (ii) *Academic/Non-profit License*, U\$ 200 a versão eletrônica e U\$ 300 a versão Web; e (iii) *Student License*, U\$ 50 a versão eletrônica e U\$ 75 a versão Web.

As Fig. 15(a) e (b) ilustram excertos do QUIS nas versão *impressa* e *HTML*, respectivamente.

5.4 Messages which appear on screen:	confusing	1	2	3	4	5	6	7	8	9	clear	NA
5.4.1 Instructions for commands or choice	confusing	1	2	3	4	5	6	7	8	9	clear	NA

(a)

Characters on the computer screen:	
hard to read	easy to read
<input type="radio"/> 1	<input type="radio"/> NA
<input type="radio"/> 2	
<input type="radio"/> 3	
<input type="radio"/> 4	
<input type="radio"/> 5	
<input type="radio"/> 6	
<input type="radio"/> 7	
<input type="radio"/> 8	
<input type="radio"/> 9	
Image of characters:	
fuzzy	sharp
<input type="radio"/> 1	<input type="radio"/> NA
<input type="radio"/> 2	
<input type="radio"/> 3	
<input type="radio"/> 4	
<input type="radio"/> 5	
<input type="radio"/> 6	
<input type="radio"/> 7	
<input type="radio"/> 8	
<input type="radio"/> 9	
Character shapes (fonts):	
barely legible	very legible
<input type="radio"/> 1	<input type="radio"/> NA
<input type="radio"/> 2	
<input type="radio"/> 3	
<input type="radio"/> 4	
<input type="radio"/> 5	
<input type="radio"/> 6	
<input type="radio"/> 7	
<input type="radio"/> 8	
<input type="radio"/> 9	
<input type="button" value="Reset"/>	

(b)

Fig. 15 - Excertos do QUIS (*Questionnaire for User Interaction Satisfaction*):
(a) versão *impressa*; e (b) versão *HTML*.

6.3.1.2 ASQ (*After-Scenario Questionnaire*), PSSUQ (*Post-Study System Usability Questionnaire*) e CSUQ (*Computer System Usability Questionnaire*)

O ASQ é um questionário de três itens que os avaliadores de usabilidade da IBM têm usado para sondar a satisfação do usuário de teste após a conclusão de cada tarefa de um roteiro de teste. Os itens são relativos a três componentes importantes de satisfação do usuário com a usabilidade do sistema: *facilidade de conclusão da tarefa*, *tempo de execução da tarefa* e *adequação das informações de suporte* (ajuda online, mensagens e documentação).

Segundo Lewis [Lew93], o ASQ apresenta a vantagem de ser curto, o que implica uma grande redução no tempo de preenchimento, uma consideração prática importante em estudos de usabilidade. Cada item do ASQ é associado a uma escala somativa ou de Likert de 7 pontos, delimitada à esquerda pelo adjetivo *strongly agree* (*concordo totalmente*) e à direita pelo adjetivo *strongly disagree* (*discordo totalmente*), conforme ilustrado na Fig. 16.

The After-Scenario Questionnaire (ASQ)

Administration and Scoring. Give the questionnaire to a participant after he or she has completed a scenario during a usability evaluation. Average (with the arithmetic mean) the scores from the three items to obtain the ASQ score for a participant's satisfaction with the system for a given scenario. Low scores are better than high scores due to the anchors used in the 7-point scales. If a participant does not answer an item or marks N/A, average the remaining items to obtain the ASQ score.

Instructions and Items. The questionnaire's instructions and items are:

For each of the statements below, circle the rating of your choice.

1. Overall, I am satisfied with the ease of completing this task.

STRONGLY
AGREE 1 2 3 4 5 6 7 **STRONGLY**
DISAGREE

2. Overall, I am satisfied with the amount of time it took to complete this task.

STRONGLY
AGREE 1 2 3 4 5 6 7 **STRONGLY**
DISAGREE

3. Overall, I am satisfied with the support information (on-line help, messages, documentation) when completing this task.

STRONGLY
AGREE 1 2 3 4 5 6 7 **STRONGLY**
DISAGREE

Fig. 16 - Estrutura do ASQ (After-Scenario Questionnaire).

Para a sondagem da satisfação do usuário com a usabilidade de sistemas de *software*, os avaliadores de usabilidade da IBM têm utilizado outro instrumento, o *PSSUQ*, composto por 19 itens, cada um dos quais associado a uma escala de Likert similar à do ASQ. Lewis [Lew93] comentou que, embora os usuários de teste gastem mais tempo para preenchê-lo (cerca de 10 minutos), o preenchem apenas uma vez, ao final do estudo de usabilidade. O preenchimento do *PSSUQ* por usuários de teste fornece uma avaliação subjetiva global do sistema que utilizaram. A Fig. 17 ilustra os três itens iniciais do *PSSUQ*.

1. Overall, I am satisfied with how easy it is to use this system.

STRONGLY
AGREE 1 2 3 4 5 6 7 **STRONGLY**
DISAGREE

COMMENTS:

2. It was simple to use this system.

STRONGLY
AGREE 1 2 3 4 5 6 7 **STRONGLY**
DISAGREE

COMMENTS:

3. I could effectively complete the tasks and scenarios using this system.

STRONGLY
AGREE 1 2 3 4 5 6 7 **STRONGLY**
DISAGREE

COMMENTS:

Fig. 17 - Excerto do PSSUQ (Post-Study System Usability Questionnaire).

Quando o estudo de usabilidade não se fundamenta em um experimento laboratorial, mas em uma avaliação de campo, os avaliadores de usabilidade da IBM substituem o *PSSUQ* por outro instrumento de sondagem da satisfação, o *CSUQ*, idêntico em estrutura ao *PSSUQ*, embora as declarações contidas nos 19 itens não se refiram a situações avaliatórias que envolvem ensaios de usabilidade. Um exemplo dessas diferenças pode ser facilmente observado a partir da comparação das Figs. 17 e 18, esta última um excerto dos três primeiros itens do *CSUQ*.

1. Overall, I am satisfied with how easy it is to use this system.								
STRONGLY								STRONGLY
AGREE	1	2	3	4	5	6	7	DISAGREE
COMMENTS:								
2. It is simple to use this system.								
STRONGLY								STRONGLY
AGREE	1	2	3	4	5	6	7	DISAGREE
COMMENTS:								
3. I can effectively complete my work using this system.								
STRONGLY								STRONGLY
AGREE	1	2	3	4	5	6	7	DISAGREE
COMMENTS:								

Fig. 18 - Excerto do *CSUQ* (*Computer System Usability Questionnaire*).

Como pôde ser observado, enquanto as duas declarações iniciais são exatamente iguais em ambos os questionários, pois se referem ao sistema como um todo, a partir do item 3 as declarações passam a diferir.

6.3.1.3 SUS (*System Usability Scale*)

O *SUS* é uma escala de usabilidade de 10 itens desenvolvida em 1986 por John Brooke [Broo96], como parte da introdução de um programa de usabilidade na *Digital Equipment Co. Ltd.* (*DEC - Reading, UK*), focalizando sistemas integrados destinados ao trabalho em escritórios.

Cada item contém uma declaração relativa ao aspecto de interesse, a qual é associada a uma escala numérica de 5 pontos (1 a 5), ladeada à esquerda pelo adjetivo *strongly disagree* (discordo totalmente) e à direita pelo adjetivo *strongly agree* (concordo totalmente), conforme ilustrado na Fig. 19.

	Strongly disagree					Strongly agree
1. I think that I would like to use this system frequently	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5	
2. I found the system unnecessarily complex	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5	
3. I thought the system was easy to use	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5	
4. I think that I would need the support of a technical person to be able to use this system	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5	
5. I found the various functions in this system were well integrated	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5	
6. I thought there was too much inconsistency in this system	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5	
7. I would imagine that most people would learn to use this system very quickly	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5	
8. I found the system very cumbersome to use	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5	
9. I felt very confident using the system	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5	
10. I needed to learn a lot of things before I could get going with this system	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
	1	2	3	4	5	

Fig. 19 - Estrutura do SUS (*System Usability Scale*).

As declarações contidas nos itens do *SUS* cobrem diversos aspectos da usabilidade do sistema avaliado, e.g. *necessidade de suporte*, *necessidade de treinamento*, *complexidade do sistema*, *consistência*, *facilidade de aprendizado* e *atitude do usuário com respeito ao sistema* (*sentimentos*).

O *SUS* é um instrumento de sondagem pós-teste, porém deve ser utilizado antes da ocorrência de qualquer discussão ou síntese pertinente ao contexto avaliatório. Segundo Brooke [Broo96], solicita-se a leitura de cada declaração e uma resposta imediata ao item, ao invés de deixar o respondente pensando sobre a declaração por muito tempo. Além disso, o avaliador recomenda ao respondente assinalar a opção central da escala quando este não se sente capaz de responder a um dado item do *SUS*.

O *SUS* produz um único valor representante de uma medida composta da usabilidade global do sistema estudado, o que significa dizer que escores para itens individuais não são *per se* significativos. O escore de cada item varia entre 0 e 4. De acordo com Brooke [Broo96], o escore do *SUS* é calculado a partir da multiplicação da soma dos escores relativos aos itens pares e ímpares. Para os itens ímpares, o escore individual é calculado subtraindo 1 da pontuação dada ao item pelo respondente, enquanto para os itens pares é calculado subtraindo de 5 a pontuação dada ao item. Os escores do *SUS* variam de 0 a 100.

O SUS tem sido disponibilizado gratuitamente para uso em estudos de usabilidade envolvendo a sondagem da opinião do usuário sobre o produto. O único pré-requisito para seu uso são os agradecimentos à fonte de mensuração, i.e., à *Digital Equipment Co. Ltd. (DEC), Reading (UK)*.

Vale a pena comentar sobre o fato de não se haver utilizado o SUS como instrumento de sondagem da opinião do usuário nesta pesquisa, uma vez que se trata de um questionário gratuito. Conforme relatado por Brooke [Broo96], o SUS seria um instrumento validado no âmbito da *Digital Equipment Co. Ltd.*, assim como em outras avaliações industriais e projetos de pesquisa. Contrariando o relato de Brooke, pesquisas realizadas no *site* da DEC e na literatura da área resultaram na constatação de que o SUS parece não ter tido muito expressividade de uso, uma vez que não há praticamente nenhuma menção de sua aplicação em estudos documentados da usabilidade de produtos.

Tendo em vista esta escala de 10 itens ter se afigurado muito superficial para a sondagem planejada no contexto desta pesquisa, além de que seria necessária a elaboração de uma versão em português, preferiu-se despende tempo e esforço na revisão e otimização dos instrumentos de sondagem já construídos em outras pesquisas do *Grupo de Interfaces Homem-Máquina (GIHM)* do DEE/CCTI/UFPB. Os dois instrumentos de sondagem resultantes dessa iniciativa serão discutidos na seção 6.3.2.

6.3.1.4 IsoMetrics

Desenvolvido por Gediga *et al.* [Gedi99] no âmbito de um projeto do *Fachbereich Psychologie* da *Universität Osnabrück*, o *IsoMetrics* é um instrumento de sondagem composto por 90 itens, cada um dos quais associado a uma escala de 5 pontos (1 a 5) de semântica diferencial, contendo as opções *stimmt nicht* (discordo totalmente), *stimmt wenig* (discordo parcialmente), *stimmt mittelmäßig* (nem discordo nem concordo), *stimmt ziemlich* (concordo parcialmente) e *stimmt sehr* (concordo totalmente). No extremo direito, integrada ao item, embora destacada da escala, há a opção *Keine Angabe* (nenhuma opinião), destinadas às circunstâncias em que o usuário se encontra indeciso, não possui uma opinião formada sobre a declaração ou não sabe o que responder.

De acordo com Gediga *et al.* [Gedi99], o processo de concepção do *IsoMetrics* teve início com a compilação de declarações presentes na literatura consagrada da área de interfaces, em guias de estilo, diretrizes de projeto, listas de inspeção, padrões e outros instrumentos de sondagem (e.g., *QUIS* [Chin88, Harp93], *Leitstand Checkliste 9241/10* [Ilg91, in Gedi99], *EVADIS I* [Oppe88, in Gedi99] e *II* [Oppe92, in Gedi99]), o que resultou em um conjunto de 651 itens. A partir de uma seleção criteriosa (e.g., eliminação de itens similares, de itens específicos a determinadas aplicações, de itens não compreensíveis por respondentes médios), foi gerado um subconjunto básico de 151 itens, que foram mais uma vez reduzidos a 90 por um grupo de especialistas em usabilidade, de companhias e universidades alemãs, com base nos sete princípios listados na Parte 10 do padrão internacional ISO 9241.

A segunda coluna do Quadro 29 discrimina o número de itens do *IsoMetrics* relacionados com cada um dos princípios do ISO 9241-10, listados na primeira coluna do quadro.

Quadro 29 – Discriminação do N° de Itens do *IsoMetrics* por Princípio do ISO 9241-10.

PRINCÍPIO DO ISO 9241-10	N° DE ITENS ASSOCIADOS
Adequação à tarefa	17
Auto-descrição	14
Facilidade de controle (<i>Controllability</i>)	14
Conformidade com as expectativas do usuário	09
Tolerância a erros	17
Adequação à individualização	11
Adequação ao aprendizado	08
Total	90

Fonte: Adaptada de Gediga et al. [Gedi99]

Segundo Gediga *et al.* [Gedi99], foram elaborados dois instrumentos de sondagem, baseados no mesmo conjunto de itens, embora usando formatos diferentes, com o propósito de atenderem às necessidades de procedimentos avaliatórios tanto somativos quanto formativos. Para avaliações somativas, o *IsoMetrics^s (short)* possui a estrutura especificada no início desta sub-seção. Por outro lado, no *IsoMetrics^t (long)*, que se destina a procedimentos avaliatórios formativos, há, além da escala de 5 pontos em que o usuário opina sobre a declaração formulada no item, uma segunda escala. Na segunda escala, com semântica diferencial distinta da primeira escala⁷⁵, o usuário declara, através do campo assinalado, a importância daquele aspecto para suas atividades.

A Fig. 20 ilustra um excerto da estrutura do *IsoMetrics^t*, enquanto a Fig. 21 ilustra os quatro primeiros itens da versão curta, o *IsoMetrics^s*. A observação mais atenta destas figuras mostra que estrutura da versão longa difere da versão curta apenas pela inclusão da segunda escala (importância subjetiva do item).

	stärkt nicht	stärkt wenig	stärkt mittelmäßig	stärkt ziemlich	stärkt sehr	Keine Angabe
A1 Die Software zwingt mich Arbeitsschritte durchzuführen, die für meine Arbeit nicht sinnvoll sind.	1	2	3	4	5	
	nicht wichtig	wenig wichtig	mittelmäßig wichtig	ziemlich wichtig	sehr wichtig	Keine Angabe
Wie wichtig ist dieser Aspekt für Ihren Gesamteindruck von der Software?	1	2	3	4	5	
Können Sie konkrete Beispiele nennen, bei denen Sie dieser Aussage zustimmen können?						

Fig. 20 - Excerto original da versão longa do *IsoMetrics 2.01*.

⁷⁵ Os adjetivos empregados no original são *nicht wichtig* (não importante), *wenig wichtig* (pouco importante), *mittelmäßig wichtig* (razoavelmente importante), *ziemlich wichtig* (muito importante) e *sehr wichtig* (muitíssimo importante).

		stimmt nicht	stimmt wenig	stimmt müdeknäsig	stimmt ziemlich	stimmt sehr	
	Aufgabenangemessenheit	1	2	3	4	5	Keine Angabe
A.1	Die Software zwingt mich Arbeitsschritte durchzuführen, die für meine Arbeit nicht sinnvoll sind.						
A.2	Die Software zwingt mich dazu einzelne Aktionen durchzuführen, die mit der eigentlichen Aufgabenbearbeitung nichts zu tun haben.						
A.3	Mit der Software kann ich ganze Arbeitsabläufe vollständig bearbeiten.						
A.4	Die Funktionen der Software unterstützen mich bei der Bearbeitung meiner Aufgaben.						

Fig. 21 - Excerto original da versão curta do *IsoMetrics 2.01*.

Na página de abertura do *Projekt IsoMetrics2* (uce psycho.uni-osnabrueck.de/isometer/) há links para páginas⁷⁶ que disponibilizam tanto a versão 2.01e (outubro 1997) do *IsoMetrics* em alemão e inglês, nos formatos *Kurzversion* (short version) e *Langversion* (long version), quanto o manual da versão 1.10 (*Das IsoMetrics – Manual 1.10*).

6.3.1.5 SUMI (Software Usability Measurement Inventory)

O *SUMI* é um questionário com 50 declarações, a cada uma das quais é associada uma escala semântica de três pontos, com as opções *agree* (concordo), *don't know* (não sei) e *disagree* (discordo). Segundo Kirakowski [Kira96], esta escala foi adotada na versão corrente do questionário porque nas versões anteriores muitos respondentes tiveram dificuldade para discriminar uma opção em gradações de cinco ou mais opções.

Elaborado por Jurek Kirakowski e tendo sido anteriormente denominado *CUSI* (*Computer User Satisfaction Inventory*), o *SUMI* é atualmente comercializado pelo *HFRG* (*Human Factors Research Group*) da *University College Cork* (Irlanda) em formato impresso (pacotes contendo 50 questionários) em 11 versões idiomáticas: inglês britânico, inglês norte-americano, holandês, sueco, norueguês, finlandês, francês, alemão, italiano, espanhol e grego. A versão *Professional*, comercializada ao preço de IR£ 955,70⁷⁷, consiste do *SUMISCO*, um aplicativo *MS Windows 3.1* e *MS Windows 95* para a computação dos escores, um manual e um pacote com 50 questionários no idioma indicado no formulário de compra.

Segundo o autor [Kira93, Kira94, Kira96], o *SUMI* é aplicável a qualquer sistema computacional que possua um terminal de vídeo, um teclado ou outro dispositivo de entrada de dados e um dispositivo periférico de armazenamento (e.g., disco rígido), também tendo sido satisfatoriamente empregado na avaliação do lado *cliente* de aplicações cliente-servidor. O

⁷⁶ uce psycho.uni-osnabrueck.de/isometer/qn.htm e uce psycho.uni-osnabrueck.de/isometer/onlinehb/homep.htm, respectivamente.

⁷⁷ E 1213,48 ou U\$ 1040,80 ou R\$ 2024,52.

tamanho mínimo do universo amostral necessário para análises com precisão aceitável a partir do seu uso é da ordem de 10 a 12 usuários, embora Kirakowski [Kira94, Kira96] mencione avaliações satisfatoriamente conduzidas com universos amostrais menores. Entretanto, o autor adverte que a generalização dos resultados do SUMI não depende tanto do tamanho de amostra em si, mas sobretudo da atenção dada ao contexto de uso do *software* e à elaboração do plano de avaliação, o qual envolve a identificação de usuários típicos do *software*, suas metas típicas e ambientes organizacional, técnico e físico nos quais o trabalho com o produto é realizado (no caso da avaliação de protótipos, isto envolve a determinação do contexto futuro de uso).

A Fig. 22(a) ilustra um excerto da parte do SUMI referente ao delineamento do perfil do respondente, enquanto a Fig. 22(b) ilustra um trecho da parte referente à sondagem da opinião do usuário de *software*.

Screening Questionnaire

Please use and adapt this questionnaire to your data gathering needs. Obviously, not all the questions are relevant or acceptable in every context! We have found it useful when preparing a questionnaire for keeping track of who our users are when we carry out a study.

User Profile

Company:.....

Job Title:.....

Age range (please check one) :

- 18 - 24
- 25 - 33
- 35 - 45
- over 45

Sex (please check one) :

- Male
- Female

Name of software being evaluated:.....

How long have you worked with this software (please check one) :

- less than 3 months
- 3 to 6 months
- 6 months to 1 year
- 2 years or more

(a)

		Disagree
		Undecided
		Agree
		↓ ↓ ↓
1	This software responds too slowly to inputs.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
2	I would recommend this software to my colleagues.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
3	The instructions and prompts are helpful.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>
4	The software has at some time stopped unexpectedly.	<input type="checkbox"/> <input type="checkbox"/> <input type="checkbox"/>

(b)

Fig. 22 - Excertos do SUMI (*Software Usability Measurement Inventory*):
(a) instrumento de delineamento do perfil do usuário; e (b) quatro itens iniciais do instrumento da sondagem da opinião do usuário.

O SUMI oferece resultados em três níveis de análise: (i) o escore *Global*, que traduz a usabilidade subjetiva global do produto (relativo a 25 itens); (ii) os escores das sub-escalas *Affect* (*Sentimento*), *Efficiency* (*Eficiência*), *Helpfulness* (*Utilidade*), *Control* (*Controle*) e *Learnability* (*Facilidade de Aprendizado*), cada uma delas associada a 10 itens; e (iii) o *ICA* (*Item Consensual Analysis*). *Affect*⁷⁸ mensura a reação emocional geral do usuário ao produto (como já fazia o *CUSI*). *Efficiency* mensura o nível de assistência do *software* ao trabalho do usuário, segundo seu ponto de vista, estando relacionado ao conceito de “transparência”. *Helpfulness*, por sua vez, mensura o grau de auto-explanação, assim como aspectos mais específicos tais como a adequação das facilidades de ajuda e da documentação. *Control* mensura o grau de controle do usuário sobre o *software*, durante a execução de suas tarefas, conforme seu sentimento. Por fim, *Learnability* mensura a rapidez e a facilidade com que o usuário é capaz de comandar o sistema ou de aprender como usar novas funcionalidades quando necessário.

Quanto ao *ICA*, trata-se de um método de análise de questionários desenvolvido, segundo Kirakowski [Kira94, Kira96], especificamente para o instrumento de sondagem *SUMI*, através do qual um banco de dados gera padrões esperados de respostas para cada item do questionário, que são comparados com os padrões reais, a partir de um teste χ^2 , gerando discrepâncias positivas ou negativas que representam, respectivamente, aspectos avaliados do sistema que podem apresentar vantagens de mercado ou aspectos que carecem ser melhor trabalhados ou otimizados.

O *HFRG* também desenvolveu e comercializa mais dois instrumentos de sondagem: o *MUMMS* (*Measuring the Usability of Multi-Media Systems*), destinado à avaliação da qualidade de uso de produtos de *software* multimídia (atualmente em desenvolvimento a versão 2.0); e o *WAMMI* (*Website Analysis and MeasureMent Inventory*), destinado à aquisição de informações sobre a usabilidade subjetiva de sites da Web e sobre a satisfação do visitante.

6.3.2 Instrumentos de Sondagem Desenvolvidos no Âmbito desta Pesquisa

Os instrumentos de sondagem desenvolvidos e administrados no contexto deste trabalho, o *DePerUSI* (*Delineamento do Perfil do Usuário de Sistemas Interativos*) e o *OpUS* (*Opinião do Usuário de Software*), tiveram como precursor um questionário pré-teste em formato convencional (lápiz e papel) elaborado e validado junto a um grupo de 60 usuários de um sistema de processamento de imagens multiespectrais⁷⁹.

Uma vez que o universo amostral de respondentes compunha-se de usuários *intermediários* e *experientes* do produto, o delineamento do perfil e a sondagem da satisfação foi feita a partir de um único instrumento de sondagem, em três níveis, a saber: (i) *O SISTEMA COMO UM TODO*, seção composta por sete itens, cujas declarações sondavam aspectos operacionais técnicos relevantes ao sistema (*funcionalidade, completeza, rapidez, confiabilidade, indulgência e documentação online e offline*); (ii) *O MODO DE COMUNICAÇÃO*, seção composta por onze itens, cujas declarações sondavam aspectos específicos a fatores humanos relevantes ao contexto

⁷⁸ Segundo Kirakowski [Kira94, Kira96], esta sub-escala pode ser encarada como *Likeability*.

⁷⁹ O *SITIM-150 v. 2.3*, desenvolvido pelo Instituto Nacional de Pesquisas Espaciais e comercializado pela Engespaço S/A [Quei94].

avaliado (universalidade da informação, usabilidade do sistema, facilidade de comunicação, motivação de uso, facilidade de memorização, retenção da atenção, volume de digitação, consistência sintática e semântica e facilidade de aprendizado); e (iii) VOCÊ E O SITIM, seção composta de 05 itens e destinada ao delineamento do perfil técnico do respondente, com base no nível acadêmico, experiência computacional prévia, frequência de uso do sistema, natureza da interação e necessidade de um operador.

A Fig. 23 ilustra um excerto do questionário que inclui o cabeçalho e as quatro primeiras declarações.

AVALIAÇÃO DA INTERFACE DE COMUNICAÇÃO COM O SITIM
Projeto de Interface Homem-Máquina - COPELE/CCTAPPs

Sua opinião sobre a comunicação homem-máquina com o SITIM é de grande importância para este trabalho. Ela será considerada para melhorar alguns aspectos relativos à sua maior satisfação como usuário. Não é necessário se identificar. Apenas crie um modo de particularizar o seu questionário, a fim de facilitar o processo de avaliação. É suficiente assinalar as respostas apresentadas (e não ser que você queira acrescentar algum outro item àqueles existentes). Portanto, por favor preencha todo o questionário.

O SISTEMA COMO UM TODO

ITEM	LEGENDA					Análise a NOTA (0 a 10) que v. atribuiria à importância do ITEM																											
	CONCORDO TOTALMENTE (CT)	CONCORDO EM MAIOR PARTE (C+)	CONCORDO EM MENOR PARTE (C-)	DISCORDO TOTALMENTE (DT)	INDECISO (I)																												
	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th colspan="5" style="text-align: center;">AVALIAÇÃO DO ITEM</th> <th colspan="6" style="text-align: center;">IMPORTÂNCIA DO ITEM</th> </tr> <tr> <th style="text-align: center;">(CT)</th> <th style="text-align: center;">(C+)</th> <th style="text-align: center;">(C-)</th> <th style="text-align: center;">(DT)</th> <th style="text-align: center;">(I)</th> <th style="text-align: center;">0</th> <th style="text-align: center;">1</th> <th style="text-align: center;">2</th> <th style="text-align: center;">3</th> <th style="text-align: center;">4</th> <th style="text-align: center;">5</th> <th style="text-align: center;">6</th> <th style="text-align: center;">7</th> <th style="text-align: center;">8</th> <th style="text-align: center;">9</th> <th style="text-align: center;">10</th> </tr> </table>					AVALIAÇÃO DO ITEM					IMPORTÂNCIA DO ITEM						(CT)	(C+)	(C-)	(DT)	(I)	0	1	2	3	4	5	6	7	8	9	10	
AVALIAÇÃO DO ITEM					IMPORTÂNCIA DO ITEM																												
(CT)	(C+)	(C-)	(DT)	(I)	0	1	2	3	4	5	6	7	8	9	10																		
01. O SITIM atende a todas as suas necessidades de processamento digital de imagens.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																	
02. O SITIM fornece respostas rápidas (modo de comunicação).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																	
03. O SITIM fornece respostas com rapidez.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																	
04. O SITIM oferece AJUDA sempre que você julgá necessária.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>																	

Fig. 23 - Excerto do cabeçalho e declarações da primeira seção (O SISTEMA COMO UM TODO) de um dos precursores dos instrumentos DePerUSI e OpUS.

A cada item do questionário foram associadas duas escalas. A primeira delas, de 5 pontos e semântica diferencial, continha as opções CT (Concordo Totalmente), C+ (Concordo em Maior Grau), C- (Concordo em Menor Grau), DT (Discordo Totalmente) e I (Indeciso), equivalentes às opções tradicionais ótimo, bom, deficiente, ruim e indeciso. Esta escala era empregada para a avaliação da declaração do item. A segunda escala, numérica de 11 pontos (0 a 10), era empregada na avaliação da importância do item para o contexto de trabalho do respondente (similarmente ao que ocorre no IsoMetrics).

A Fig. 24 destaca um excerto da seção de delineamento do perfil técnico do respondente,

a qual continha itens destinados à sondagem de características *relativas ao conhecimento e à experiência* (e.g., nível acadêmico, tempo de uso) e à *tarefa e ao trabalho* (e.g., natureza das atividades desenvolvidas com o auxílio do produto).

VOCÊ E O SITIM	
<p>01. Você é:</p> <p><input type="checkbox"/> Estudante de Graduação</p> <p><input type="checkbox"/> Graduado</p> <p><input type="checkbox"/> Pós-Graduado</p>	<p>05. A natureza das atividades que você desenvolve com o auxílio do SITIM é de:</p> <p><input type="checkbox"/> Pesquisa e desenvolvimento</p> <p><input type="checkbox"/> Aplicação imediata de usuários</p> <p><input type="checkbox"/> Treinamento e aprendizagem</p>
<p>02. De que equipamento você dispõe:</p> <p><input type="checkbox"/> SITIM-150</p> <p><input type="checkbox"/> SITIM-340</p>	<p>07. Você já trabalhou com outro sistema de processamento de imagens além do SITIM?</p> <p><input type="checkbox"/> Sim</p> <p><input type="checkbox"/> Não</p>
<p>03. Com que versão do SITIM você trabalha?</p> <p><input type="checkbox"/> Versão 2.1</p> <p><input type="checkbox"/> Versão ??</p>	<p>08. Há quanto tempo você utiliza processamento digital de imagens como ferramenta em sua área de atuação?</p>

Fig. 24 - Excerto seção referente ao delineamento do perfil técnico do respondente (*VOCÊ E O SITIM*) de um dos precursores dos instrumentos *DePerUSI* e *OpUS*.

Em 1998, Queiroz e Turnell [Quei98a, Quei98b] reestruturaram a seção *VOCÊ E O SITIM*, com o propósito de convertê-la em um instrumento mais genérico de sondagem do perfil do usuário de *software*, a ser administrado antes da realização de teste de usabilidade de produtos de interesse. O instrumento de delineamento do perfil do usuário passou a ser denominado *questionário pré-teste*, tendo sido administrado entre as etapas de recrutamento de participantes e de condução do ensaio de usabilidade destinado à avaliação do módulo de entrada de um SIG.

O referido questionário, resultado da reestruturação da seção de sondagem do perfil do usuário do questionário desenvolvido por Queiroz [Quei94], deixou de ser específico a um produto, tendo sido anexado a um documento que detalhava, para o participante potencial do ensaio de usabilidade do SIG, todas as etapas do processo de avaliação planejado, o qual consistia (i) do preenchimento de um questionário pré-teste (delineamento do perfil do participante), (ii) da condução de testes laboratoriais; e (iii) do preenchimento de um questionário pós-teste (sondagem da opinião do participante). A declaração das condições de teste também especificava o produto a ser avaliado e os direitos do participante.

A Fig. 25 apresenta um excerto do *questionário pré-teste* acima descrito, que continha 14 itens, 11 de múltipla escolha e 3 de preenchimento com texto, o qual foi administrado no formato impresso.

QUESTIONÁRIO PRÉ-TESTE

Este questionário tem o propósito de coletar informações sobre você enquanto usuário de sistemas computacionais e, em particular, do PRODUTO sob condições de teste. Por favor assinale um X na opção que mais se adequa à sua opinião sobre cada aspecto considerado e forneça as demais informações solicitadas, quando se fizer necessário.

VOCÊ E O PRODUTO			
Você é	Pós-graduado <input type="checkbox"/>	Graduado <input type="checkbox"/>	Estudante de graduação <input type="checkbox"/>
A natureza das atividades que v. desenvolve com o auxílio do PRODUTO é de:	Pesquisa & Desenvolvimento <input type="checkbox"/>	Aplicação imediata de utilitários <input type="checkbox"/>	Treinamento & Aprendizagem <input type="checkbox"/>
Você tem experiência prévia com sistemas computacionais?	SIM <input type="checkbox"/> NÃO <input type="checkbox"/>		
Há quanto tempo v. usa sistemas computacionais?	Menos de 3 meses <input type="checkbox"/>	3 meses a 1 ano <input type="checkbox"/>	Mais de 1 ano <input type="checkbox"/>
Você tem experiência prévia com o PRODUTO?	SIM <input type="checkbox"/> NÃO <input type="checkbox"/>		

Fig. 25 - Excerto do questionário pré-teste, uma reestruturação do questionário elaborado por Queiroz [Quei94], feita por Queiroz e Turnell [Quei98a, Quei98b].

Paralelamente à elaboração do *questionário pré-teste*, foi desenvolvido, a partir dos itens contidos nas duas primeiras seções do questionário de Queiroz [Quei94], um instrumento complementar para a sondagem da satisfação subjetiva do usuário de *software*. Este instrumento foi denominado *questionário pós-teste*, devendo ser preenchido pelo usuário de teste imediatamente após a execução da última tarefa.

O propósito da administração do *questionário pós-teste* no contexto da avaliação realizada por Queiroz e Turnell [Quei98a, Quei98b] diferia daquele descrito por Queiroz [Quei94] principalmente no que diz respeito à natureza dos dois procedimentos de avaliação. Enquanto este último era de natureza formativa (avaliação voltada para o reprojeto da interface), o primeiro era eminentemente somativo (avaliação destinada à emissão de um parecer sobre a interface do produto). Eis porque a escala de importância subjetiva do item foi omitida no *questionário pós-teste*.

O *questionário pós-teste* cumpria-se de três seções, a saber: (i) *SOFTWARE*, com 10 itens, que focalizava aspectos da interface relativos ao uso e à navegação; (ii) *DOCUMENTAÇÃO ONLINE E OFFLINE*, com 08 itens relativos à ajuda oferecida pelo produto, tanto em nível do manual do usuário quanto em termos dos mecanismos disponibilizados *online*; e (iii) *ASPECTOS ADICIONAIS*, com 07 itens em aberto, a serem completados pelo usuário, conforme sua visão da abrangência da abordagem das seções anteriores.

Cada um dos 25 itens do *questionário pós-teste* foi acompanhado de uma escala de 5 pontos (1 a 5) de semântica diferencial, contendo os adjetivos *muito fácil* (1), *fácil* (2), *nem fácil nem difícil* (3), *difícil* (4) e *muito difícil* (5). A cada item também foi adicionado um campo para a inclusão de comentários que o usuário achasse pertinentes ao contexto do item. A Fig. 26

apresenta um excerto do *questionário pós-teste* que, assim como o *questionário pré-teste*, foi administrado no formato *impresso*.

QUESTIONÁRIO PÓS-TESTE

Este questionário integrado tem o propósito de coletar informações sobre como você se sente a respeito do uso do software e da documentação *online* e *offline* utilizada na(s) sessão(ões) de teste. Por favor assinale um X no item da escala que mais se adequa à sua opinião sobre cada aspecto do produto considerado neste questionário.

<i>Software</i>					
ASPECTO	ESCALA				
	1	2	3	4	5
	Muito fácil	Fácil	Nem fácil nem difícil	Difícil	Muito difícil
1. Uso do software na realização de tarefas de interesse					
	Comentário				
2. Localização dos itens de menu associados às tarefas					
	Comentário				
3. Comunicação com o software (terminologia, linguagem, retorno da informação e das ações em geral)					
	Comentário				

Fig. 26 - Excerto do *questionário pós-teste*, uma reestruturação do questionário elaborado por Queiroz [Quei94], feita por Queiroz e Turnell [Quei98a, Quei98b].

6.3.2.1 DePerUSI (*Delineamento do Perfil do Usuário de Sistemas Interativos*)

O *DePerUSI* é um instrumento de sondagem destinado à coleta de informações que auxiliem o avaliador no delineamento do perfil dos usuários de teste de ensaios de usabilidade fundamentados em produtos de aplicação genérica ou específica os mais diversos, e.g., editores de texto, planilhas eletrônicas, editores de páginas para a Web, sistemas de processamento de imagens, sistemas de informações geográficas.

O *DePerUSI* contém 24 itens, destinados à sondagem de 19 características (i) *físicas*, (ii) *relativas ao conhecimento e à experiência*; e (iii) *relativas à tarefa e ao trabalho*, discriminadas no Quadro 30. Estas características possibilitam o “mapeamento” de universos amostrais de usuários envolvidos em diferentes contextos avaliatórios, facilitando sua categorização (e.g., principiantes, intermediários e experientes ou esporádicos e freqüentes).

Quadro 30 – Características relativas ao usuário incluídas no *DePerUSI*.

NATUREZA DA CARACTERÍSTICA	ESPECIFICAÇÃO
Física	Sexo
	Destreza Manual
	Uso de Corretivos Visuais
	Faixa Etária
Relativa ao Conhecimento e à Experiência	Nível Acadêmico
	Experiência Computacional Prévia
	Tempo de Uso de Sistemas Computacionais
	Frequência de Uso de Sistemas Computacionais
	Experiência Prévia com o Produto
	Tempo de Uso do Produto
	Frequência de Uso do Produto
	Experiência Prévia com Produtos Similares
	Discriminação de Produtos Similares Utilizados
Tempo de Uso de Produtos Similares	
Relativa à Tarefa e ao Trabalho	Natureza das Atividades (utilizando o produto)
	Uso Contextual do Produto
	Plataforma Computacional
	Versão Utilizada do Produto
	Forma de Ajuda mais Utilizada

Características físicas, tais como *destreza manual* e *uso de corretivos visuais*, podem ser úteis em inferências sobre problemas relativos ao ambiente físico (e.g., adequação dos dispositivos de entrada à destreza manual do usuário, adequação das condições de iluminação aos usuários portadores de corretivos visuais).

Adicionalmente, características relativas ao trabalho, de um modo geral, assim como a tarefa desenvolvida com o auxílio do produto avaliado, tais como *uso contextual* e *forma de ajuda mais utilizada* podem oferecer ao avaliador informações que, integradas a dados de outras naturezas (e.g., cálculos e medições), o auxiliem a propor recomendações que possibilitem a otimização do produto em um dado contexto de uso, assim como dos recursos de assistência ao usuário oferecidos pelo desenvolvedor.

Inicialmente desenvolvido em formato eletrônico, integrando recursos do *Visual Basic* às facilidades de processamento matemático e estatístico de dados oferecidas pelo *MS Excel 2000*, o *DePerUSI.xls* assumiu, logo em seguida, outro formato eletrônico, ao ter a estrutura convertida para o formato de documento do *MS Word 2000*.

Vale a pena salientar que, apesar de mais flexível em termos de adaptação dos itens a novos contextos avaliatórios, a versão *DePerUSI.doc* não possui a versatilidade de totalização dos dados apresentada pela versão *.xls*, embora facilite a geração de cópias impressas do *DePerUSI*.

A idéia inicial era desenvolver um instrumento de sondagem que possibilitasse uma distribuição mais efetiva junto aos usuários fisicamente menos acessíveis, assim como um retorno mais rápido das respostas fornecidas pelo usuário, aliado a um processo de compilação dos dados mais rápido e fácil.

Acreditava-se *a priori* que o desenvolvimento do questionário a partir de aplicações do *MS Office*, aliado a sua distribuição e recebimento via correio eletrônico, asseguraria a efetividade e

rapidez do processo de coleta de dados relativos à sondagem via questionário, tendo em vista ter-se constatado que no âmbito do universo amostral de usuários de teste potenciais – a UFPB, a maioria dos sistemas computacionais possuía o *MS Office* instalado, situação via de regra reproduzida nos ambientes dos prováveis respondentes. Entretanto, a primeira dificuldade surgiu logo no início do processo de distribuição do *DePerUSI*, quando diversos usuários alegaram não ter acesso fácil a recursos de correio eletrônico. Tal fato implicou a necessidade de geração de um formato impresso para o questionário. Assim, surgiram as versões eletrônica *.doc* e *impressa*.




Grupo de Interfaces Homem-Máquina



DELINEAMENTO DO PERFIL DO USUÁRIO DE SISTEMAS INTERATIVOS

Este questionário tem o propósito de coletar informações que possibilitem delinear seu perfil de usuário de sistemas computacionais e, em particular, do PRODUTO sob condições de teste. Por favor analise cada aspecto considerado, selecionando a opção que mais adequa à sua condição de usuário e fornecendo as demais informações solicitadas, quando se fizer necessário.

Qualquer dúvida no processo de preenchimento do questionário, consultar o suporte técnico nos telefones (021)93 310 1140 ou (021)93 310 1357.

ITEM	OPÇÕES			
1. Você é:	<input type="checkbox"/> PÓS-GRADUADO	<input type="checkbox"/> GRADUADO	<input type="checkbox"/> ESTUDANTE DE GRADUAÇÃO	<input type="checkbox"/> ESTUDANTE DE PÓS-GRADUAÇÃO
2. Você é do sexo:	<input type="checkbox"/> MASCULINO	<input type="checkbox"/> FEMININO		
3. Você é:	<input type="checkbox"/> DESTRO	<input type="checkbox"/> CANHOTO	<input type="checkbox"/> AMBIDESTRO	
4. Você usa óculos ou lentes de contato?	<input type="checkbox"/> SIM	<input type="checkbox"/> NÃO		
5. Você pertence à faixa etária de:	<input type="checkbox"/> 18 A 24 ANOS	<input type="checkbox"/> 25 A 35 ANOS	<input type="checkbox"/> 35 A 45 ANOS	<input type="checkbox"/> ACIMA DE 45 ANOS
6. A natureza da principal atividade que v. desenvolve com o auxílio do PRODUTO é essencialmente de:	<input type="checkbox"/> PESQUISA E DESENVOLVIMENTO	<input type="checkbox"/> APLICAÇÃO IMEDIATA DE UTILITÁRIOS	<input type="checkbox"/> TREINAMENTO E APRENDIZAGEM	
7. No contexto de suas atividades v. utiliza o PRODUTO em:	<input type="checkbox"/> Construção de Modelos <input type="checkbox"/> Prestação de Serviços <input type="checkbox"/> Suporte a Disciplinas Teóricas (Graduação)	<input type="checkbox"/> Verificação de Modelos Existentes <input type="checkbox"/> Execução de Projetos <input type="checkbox"/> Suporte a Disciplinas Teóricas (Pós-graduação)	<input type="checkbox"/> Utilização em Teses/Dissertações <input type="checkbox"/> Desenvolvimento de Produtos <input type="checkbox"/> Suporte Laboratorial (Graduação)	<input type="checkbox"/> Uso em Trabalhos de Iniciação Científica <input type="checkbox"/> Treinamento em Nível de Extensão <input type="checkbox"/> Suporte Laboratorial (Pós-graduação)
8. Sua plataforma computacional é:	<input type="checkbox"/> UNIX	<input type="checkbox"/> PC	<input type="checkbox"/> MAC	<input type="checkbox"/> OUTRA
9. Você tem experiência prévia com sistemas computacionais?	<input type="checkbox"/> SIM <input type="checkbox"/> NÃO			
10. Há quanto tempo v. usa sistemas computacionais?	<input type="checkbox"/> MENOS DE 3 MESES	<input type="checkbox"/> DE 3 MESES A 1 ANO	<input type="checkbox"/> MAIS DE 1 ANO	
11. Com que frequência v. usa sistemas computacionais?	<input type="checkbox"/> DIARIAMENTE <input type="checkbox"/> 1 VEZ POR QUINZANA	<input type="checkbox"/> MENOS DE 1 VEZ POR SEMANA <input type="checkbox"/> MENOS DE 1 VEZ POR QUINZANA	<input type="checkbox"/> 1 VEZ POR SEMANA <input type="checkbox"/> 1 VEZ POR MÊS	<input type="checkbox"/> MENOS DE 1 VEZ POR SEMANA <input type="checkbox"/> É A 1ª VEZ QUE USO O PRODUTO

(a)

Grupo de Interfaces Homem-Máquina

DELINEAMENTO DO PERFIL DO USUÁRIO DE SISTEMAS INTERATIVOS

Este questionário tem o propósito de coletar informações que possibilitem delinear seu perfil de usuário de sistemas computacionais e, em particular, do PRODUTO sob condições de teste. Por favor analise cada aspecto considerado, selecionando a opção que mais adequa à sua condição de usuário e fornecendo as demais informações solicitadas, quando se fizer necessário.

ASPECTO	OPÇÕES			
1. Você é:	NÍVEL ACADÊMICO <input type="radio"/> PÓS-GRADUADO <input type="radio"/> GRADUADO <input type="radio"/> ESTUDANTE DE PÓS-GRADUAÇÃO <input type="radio"/> ESTUDANTE DE GRADUAÇÃO			
2. Você é do sexo:	SEXO <input type="radio"/> MASCULINO <input type="radio"/> FEMININO			
3. Você é:	DESTREZA MANUAL <input type="radio"/> DESTRO <input type="radio"/> CANHOTO <input type="radio"/> AMBIDESTRO			
4. Você usa óculos ou lentes de contato?	USO DE CORRETIVOS VISUAIS <input type="radio"/> SIM <input type="radio"/> NÃO			
5. Você pertence à faixa etária de:	FAIXA ETÁRIA <input type="radio"/> 18 A 24 ANOS <input type="radio"/> 25 A 35 ANOS <input type="radio"/> 35 A 45 ANOS <input type="radio"/> ACIMA DE 45 ANOS			
6. A natureza da principal atividade que v. desenvolve com o auxílio do PRODUTO é essencialmente de:	NATUREZA DAS ATIVIDADES COM O PRODUTO <input type="radio"/> PESQUISA E DESENVOLVIMENTO <input type="radio"/> APLICAÇÃO IMEDIATA DE UTILITÁRIOS <input type="radio"/> TREINAMENTO E APRENDIZAGEM			
7. No contexto de suas atividades v. utiliza o PRODUTO em:	<input type="checkbox"/> Construção de Modelos <input type="checkbox"/> Prestação de Serviços <input type="checkbox"/> Suporte a Disciplinas Teóricas (Graduação)	<input type="checkbox"/> Verificação de Modelos Existentes <input type="checkbox"/> Execução de Projetos <input type="checkbox"/> Suporte a Disciplinas Teóricas (Pós-graduação)	<input type="checkbox"/> Utilização em Teses/Dissertações <input type="checkbox"/> Desenvolvimento de Produtos <input type="checkbox"/> Suporte Laboratorial (Graduação)	<input type="checkbox"/> Uso em Trabalhos de Iniciação Científica <input type="checkbox"/> Treinamento em Nível de Extensão <input type="checkbox"/> Suporte Laboratorial (Pós-graduação)
8. Sua plataforma computacional é:	PLATAFORMA COMPUTACIONAL <input type="radio"/> UNIX <input type="radio"/> PC <input type="radio"/> MAC <input type="radio"/> OUTRA			
9. Você tem experiência prévia com sistemas computacionais?	EXPERIÊNCIA COMPUTACIONAL PRÉVIA <input type="radio"/> SIM <input type="radio"/> NÃO			
10. Há quanto tempo v. usa sistemas	TEMPO DE USO DE SISTEMAS COMPUTACIONAIS			

(b)

Fig. 27 - Excertos do *DePerUSI*: (a) versão *impressa*; e (b) versão *.xls*.

A necessidade de coleta do maior número possível de instrumentos de sondagem, a fim de assegurar a significância dos resultados obtidos a partir desta estratégia, implicou o envio do *DePerUSI* para usuários do *MATLAB* atuantes em outras instituições. A segunda dificuldade surgiu quando vários desses prováveis respondentes alegaram não dispor do *MS Office* em seus sistemas computacionais, o que impediu a visualização e o preenchimento dos questionários enviados. Esta dificuldade foi decisiva para a geração da versão *DePerUSI.html*. A partir dela, qualquer usuário do *MATLAB* pode acessar o questionário, preenchê-lo e devolvê-lo via Web.

As Figs. 27(a) e (b) ilustram, respectivamente, excertos das versões *DePerUSI impresso* e *DePerUSI.xls*, enquanto a versão *DePerUSI.html* encontra-se disponível no endereço <http://www.dee.ufpb.br/~copele/qihm/index.html>. Por sua vez, o Anexo G apresenta uma cópia da tela da versão *.xls* do *DePerUSI*.

6.3.2.2 OpUS (*Opinião do Usuário de Software*)

O *OpUS* é um instrumento destinado à sondagem da opinião do usuário de sistemas computacionais sobre diversos aspectos de um dado produto de *software* de aplicação genérica ou específica (e.g., editores de texto, planilhas eletrônicas, sistemas de processamento de imagens, sistemas de informações geográficas) focalizado por avaliadores de interfaces como alvo de estudo da usabilidade.

Conforme anteriormente mencionado, o *OpUS* teve o mesmo precursor do *DePerUSI*, i.e. o questionário desenvolvido por Queiroz [Quei94] para a sondagem da satisfação subjetiva de usuários de um sistema de processamento de imagens multiespectrais. Dado o redirecionamento dos propósitos avaliatórios, o questionário de Queiroz [Quei94] foi reestruturado em dois instrumentos de sondagem (*questionários pré-teste* e *pós-teste*). Sua administração, intercalada pelo ensaio de usabilidade do estudo de caso feito por Queiroz e Turnell [Quei98a, Quei98b], forneceu, além das informações almejadas, diversos indicativos para sua reestruturação, resultando nos dois instrumentos de sondagem utilizados nesta pesquisa.

Em nível da administração do *questionário pós-teste*, foram verificados dois eventos que se mostraram relevantes para a concepção do *OpUS*. O primeiro deles foi o não preenchimento dos campos de comentários pela quase totalidade dos respondentes. Tal ocorrência fez com que o campo de comentários fosse omitido do novo instrumento de sondagem.

O segundo evento, observado durante o preenchimento do *questionário pós-teste*, foi a pouca importância que cerca de 60% dos respondentes deram aos itens em aberto. Alguns deles chegaram a comentar que não conseguiam formular declarações adicionais, sugerindo um re-estudo das duas primeiras seções e uma complementação da seção *ASPECTOS ADICIONAIS* com declarações que pudessem fornecer informações incrementais de interesse do avaliador.

Assim, o *questionário pós-teste* foi revisado, a partir de uma comparação com os instrumentos de sondagem descritos na sub-seção 6.3.1, tendo sido reformulado de modo a manter a estrutura de 03 seções, embora acrescidas de novos itens.

As 03 seções do *OpUS*, conforme se pode observar no Anexo H, são: (i) *O Produto – Uso e Navegação*, contendo 20 itens e visando a aquisição de informações de natureza similar àquela da seção *SOFTWARE* do *questionário pós-teste*; (ii) *O Produto – Documentação Online e Offline*, contendo 10 itens referentes a aspectos similares àqueles sondados através da seção *DOCUMENTAÇÃO ONLINE E OFFLINE* do *questionário pós-teste*; e (iii) *Você e o Produto*, composta por 20 itens referentes ao “sentimento” do respondente sobre nuances de alguns dos aspectos sondados nas seções precedentes, assim como sobre aspectos relacionados com sua aceitação do produto.

As 30 primeiras declarações do *OpUS* são associadas à mesma escala de 5 pontos de semântica diferencial adotada no *questionário pós-teste*, constituída dos adjetivos *muito fácil* (1), *fácil* (2), *nem fácil nem difícil* (3), *difícil* (4) e *muito difícil* (5). No entanto, dada a natureza e o estilo das 20 últimas declarações, a escala semântica das seções anteriores foi substituída por outra escala contendo os adjetivos *concordo totalmente* (1), *concordo parcialmente* (2), *nem concordo nem discordo* (3), *discordo parcialmente* (4) e *discordo totalmente* (5).

Similarmente ao *DePerUSI*, o *OpUS* pode ser administrado nas versões *XLS*, *impressa* e *HTML*. As Figs. 28 e 29 ilustram, respectivamente, excertos da primeira e última seções do *OpUS.xls*, onde são destacadas as escalas semânticas comentadas no parágrafo anterior.

A versão *OpUS.html* pode ser acessada a partir do endereço <http://www.dee.ufpb.br/~copele/gihm/index.html>.

Este questionário integrado tem o propósito de coletar informações acerca de como você se sente a respeito do uso do PRODUTO sob condições de teste e de sua documentação *online* e *offline*. Por favor analise cada aspecto considerado, selecionando a opção que mais adequa à sua condição de usuário.

ASPECTO		ESCALA				
		1	2	3	4	5
		Muito fácil	Fácil	Nem fácil nem difícil	Difícil	Muito difícil
1	Uso do produto na realização de tarefas de Interesse	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2	Comunicação com o produto (terminologia, linguagem, retorno da informação e das ações em geral)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3	Localização dos itens de menu associados às tarefas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4	Acesso aos itens de menu associados às tarefas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Escala numérico-semântica

Fig. 28 - Excerto da seção inicial do instrumento desenvolvido para a sondagem da *Opinião do Usuário de Software* (versão *OpUS.xls*).

Você e o Produto		ESCALA				
ASPECTO		1	2	3	4	5
		Concordo totalmente	Concordo parcialmente	Nem concordo nem discordo	Discordo parcialmente	Discordo totalmente
31	Acho que a apresentação do produto é bastante atraente, o que estimula seu uso por mim.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32	Acho que a apresentação do produto facilita o aprendizado rápido de seus comandos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
33	Acho o modo de apresentação das mensagens do produto bastante <u>claro</u> e <u>compreensível</u> .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
34	A resposta do produto às minhas entradas é muito lenta.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
35	Sempre me sinto no controle das ações quando uso o produto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
36	Consigo executar as tarefas de modo <u>direto</u> ao usar o produto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
37	Acho que o produto atende plenamente às minhas necessidades.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
38	Perco muito tempo tentando aprender os comandos do produto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
39	Consigo fazer <u>exatamente</u> o que desejo com os recursos oferecidos pelo produto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Fig. 28 - Excerto da seção final do instrumento desenvolvido para a sondagem da *Opinião do Usuário de Software* (versão *OpUS.xls*).

6.4 Metodologia Adotada

Conforme explicitado no capítulo introdutório, o propósito essencial desta pesquisa é a confrontação dos enfoques avaliatórios, centrados na usabilidade de produtos interativos usuário-computador – *mensuração do desempenho*, *mensuração da satisfação* e *inspeção de padrões* – visando a investigação da influência do atributo central⁸⁰ de cada enfoque considerado sobre a usabilidade da interface de um aplicativo de *software*.

Com base nas metas e interesses inseridos neste propósito e traduzidos pelos objetivos discriminados no primeiro capítulo, foi realizada uma sondagem de campo junto aos laboratórios do *Departamento de Engenharia Elétrica (UFPB/CCT/DEE)*, com o propósito de verificar os aplicativos de uso específico utilizados por um contingente de professores e alunos de graduação e pós-graduação que fosse, além de representativo, acessível ao recrutamento para a participação não remunerada em um ensaio de usabilidade. Tal sondagem possibilitou não somente a caracterização do alvo de estudo, como também a caracterização do universo amostral, discriminado em três categorias de usuários de teste – *principiantes*, *intermediários* e *experientes*.

A primeira escolha recaiu sobre o *Ptolemy v. 0.6 (The University of California)*, um aplicativo destinado ao projeto e simulação de sistemas heterogêneos, empregando diferentes modelos computacionais implementados em domínios distintos. Entretanto, a partir do levantamento dos usuários deste aplicativo no âmbito do *Departamento de Engenharia Elétrica*, constatou-se a insuficiência do número de usuários de teste potenciais pertencentes às categorias previamente discriminadas.

⁸⁰ *Desempenho* para a *mensuração do desempenho*, *satisfação subjetiva do usuário* para a *mensuração da satisfação* e *grau de adequação a um padrão* para o enfoque *inspeção de padrões*.

Como alternativa, a escolha recaiu sobre o *MATLAB v. 5.3.0 (The MathWorks Inc.)*, um ambiente de computação de alto desempenho, destinado ao processamento numérico e à visualização de dados relativos a diversos problemas técnicos cujas simulações são traduzidas em formulações vetoriais e matriciais.

A concepção do procedimento metodológico relativo à avaliação do *MATLAB v. 5.3.0* fundamentou-se tanto nos objetivos a serem atingidos quanto nas condições materiais, humanas e econômicas restritivas (*guerrilla HCI [Niel94b, Hell94]*) do *Laboratório de Interfaces Homem-Máquina (LIHM)* do *DEE/CCT/UFPB*, i.e. *número reduzido de avaliadores e de recursos instrumentais, recursos financeiros limitados e número reduzido de usuários de teste disponíveis para a participação não remunerada*).

A metodologia adotada nesta pesquisa é, assim como o são os instrumentos de sondagem descritos na seção anterior, fundamentalmente uma reestruturação da metodologia concebida e validada por Queiroz e Turnell [Quei98a, Quei98b]. A diferença mais significativa no estudo revisivo desta metodologia é a inclusão do processo de inspeção de conformidade do produto com as *Partes 14, 15 e 16* do padrão internacional *ISO 9241*, não apenas como uma técnica adicional de avaliação, mas como um enfoque avaliatório complementar de expressividade sobretudo no que concerne ao diagnóstico de produtos comercializados em nível internacional.

As sub-seções seguintes descrevem, de forma detalhada ou sumariada, conforme a pertinência ao contexto, as etapas que compuseram o procedimento metodológico utilizado nesta pesquisa.

6.4.1 Planejamento do Ensaio Avaliatório

O planejamento de qualquer ensaio avaliatório demanda usualmente a caracterização do alvo de estudo, a definição de metas de interesse, a caracterização do universo amostral a ser considerado, o levantamento de indivíduos que apresentem o potencial para o recrutamento e seleção como usuários de teste, a definição do modo de recrutamento, a triagem dos usuários selecionados para a definição do número de participantes, a seleção das técnicas de avaliação mais adequadas às condições materiais, humanas e econômicas vigentes e a definição de indicadores objetivos (quantitativos) e subjetivos (qualitativos) a serem considerados.

Cada uma destas iniciativas pode ser encarada como uma sub-etapa do planejamento do ensaio, pois facilita a compartimentação das atividades associadas a cada uma delas e, por conseguinte a sistematização do processo. É deste modo que serão tratadas nesta sub-seção.

6.4.1.1 Caracterização do Alvo de Estudo

Esta sub-etapa caracterizou-se pela definição do produto-alvo da pesquisa. Em situações reais, a avaliação de um produto final usualmente ocorre com fins de otimização do produto ou de solução de um problema ou de uma série de problemas detectados pelos seus desenvolvedores ou usuários. Nesta pesquisa, a seleção do produto-alvo do estudo foi função da diversidade de categorias de usuários (experientes, intermediários e principiantes) disponíveis para fins de recrutamento como participantes dos ensaios avaliatórios.

6.4.1.2 Definição das Metas e Interesses

Após a caracterização do alvo de estudo a ser considerado nesta pesquisa, foram especificadas as metas e interesses genéricos e específicos que pautaram a condução do processo avaliatório através do qual serão verificadas as hipóteses formuladas. As metas e interesses desta pesquisa são traduzidas pelos objetivos já explicitados, devendo o grau de detalhamento ser aprofundado ao longo da discussão dos resultados obtidos.

6.4.1.3 Caracterização do Universo Amostral

Durante esta sub-étapa, estabeleceu-se, com base nas metas e interesses especificados, as características relevantes para o delineamento dos perfis dos usuários de teste. Tal definição possibilitou tanto o planejamento do modo de partição do universo amostral em sub-grupos de usuários (e.g., experientes, intermediários e principiantes) quanto a submissão dos dados coletados a um tratamento estatístico mais acurado.

Tendo em vista não se tratar pura e simplesmente do processo avaliatório de um produto interativo, mas da investigação científica de compromissos entre atributos de usabilidade em nível de desenvolvimento e avaliação de produtos interativos, caracterizou-se o universo amostral a partir da escolha de um produto de teste empregado por indivíduos com perfis afins (professores e estudantes de graduação e pós-graduação) para a solução de diferentes níveis de problemas, o que possibilitou a discriminação das três categorias de usuários de teste necessárias - *experientes*, *intermediários* e *principiantes*.

6.4.1.4 Levantamento dos Usuários de Teste Potenciais

A caracterização do universo amostral almejado conduz ao processo de mapeamento do contingente de usuários de teste potenciais e à sondagem informal da disponibilidade e interesse na participação nos ensaios pertinentes ao contexto da pesquisa.

A partir do levantamento preliminar do universo amostral potencial considerado no contexto desta pesquisa (286 indivíduos), foi possível selecionar e recrutar 58 usuários de teste, embora se houvesse intentado, desde o início do planejamento, trabalhar com um universo amostral de, pelo menos, 60 usuários. Como se costuma usualmente recrutar um contingente superior àquele com o qual se deseja trabalhar, a fim de compensar desistências e desvios significativos nos resultados, teria sido necessário o recrutamento de, pelo menos, 70 usuários para uma margem de exclusão próxima de 10% do montante de usuários recrutados.

6.4.1.5 Definição do Modo de Recrutamento dos Participantes

Esta sub-étapa caracteriza-se usualmente pela definição e estruturação de uma estratégia a ser adotada no processo de recrutamento dos participantes (e.g., envio de correspondência, aplicação de questionário pré-teste, contato telefônico), através da qual a sondagem da disponibilidade e interesse dos usuários potenciais do produto-alvo será formalizada.

Como já foi ressaltado ao longo deste documento, durante esta pesquisa, Queiroz e Turnell [Quei98a, Quei98b] desenvolveram um procedimento metodológico contextual, direcionado para a

avaliação do módulo de entrada de dados de um SIG e fundamentado na mensuração do desempenho e da satisfação subjetiva do universo amostral considerado. Tal procedimento foi concebido *a priori* com o propósito de investigação da eficácia e eficiência das técnicas adotadas para o contexto considerado, embora posteriormente se tenha concentrado os esforços na generalização de suas etapas, visando torná-lo flexível e adaptável a outros contextos avaliatórios.

Além de fundamentar a condução de ensaios avaliatórios em outros contextos de trabalho, e.g. *sites* da Web destinados ao fornecimento de informações georreferenciadas [Alme00] e aplicações de banco de dados [Niga00], o procedimento metodológico concebido e o material associado (e.g., declaração das condições de teste, instrumentos de sondagem, roteiros de tarefas) embasaram esta pesquisa de diversos modos, como será discutido a seguir. No tocante ao recrutamento de participantes, o material utilizado neste ensaio é fruto da revisão e aprimoramento daquele anteriormente elaborado por Queiroz e Turnell [Quei98a, Quei98b].

6.4.1.6 Decisão do Número de Participantes do Ensaio Avaliatório

Nesta sub-etapa, os resultados da sondagem formal do universo potencial de usuários de teste foram confrontados com os resultados do planejamento dos sub-grupos de teste, respaldando a tomada de decisão relativa ao número de participantes do ensaio, assim como a partição efetiva desse universo amostral em sub-grupos de teste, conforme será constatado no próximo capítulo.

A decisão tomada no tocante ao número de usuários de teste dos ensaios de usabilidade considerou um universo amostral de 58 participantes, dividido em três sub-grupos, conforme a experiência com o produto e a frequência de uso. Entretanto, apenas cerca de 74% do contingente considerado, ou seja 43 indivíduos, participaram das sessões de testes destinadas à avaliação do desempenho e da satisfação dos usuários com o produto-alvo. Deste montante, os resultados obtidos nos testes de aproximadamente 5% dos participantes apresentaram desvios significativos do restante, tendo sido excluídos do universo amostral.

Assim, o universo amostral efetivo deste experimento totalizou 40 indivíduos, de ambos os sexos, com predominância do sexo masculino, distribuídos em três categorias - *principiantes*, *intermediários* e *experientes*, na proporção de 16:16:8, respectivamente.

6.4.1.7 Seleção das Técnicas de Avaliação

Conforme mencionado previamente, as técnicas de avaliação da usabilidade do produto-alvo, fundamentadas na mensuração do desempenho e da satisfação subjetiva do usuário, foram definidas tanto em função dos recursos humanos, físicos e materiais disponíveis quanto de todas as informações coletadas ao longo das sub-etapas anteriores e dos estudos realizados.

Além do procedimento metodológico contextual concebido, também foi desenvolvido um estudo revisivo [Quei96] da literatura, a partir do qual as técnicas de avaliação de produtos interativos homem-máquina foram categorizadas em: (i) *ensaios de usabilidade*, onde foram discutidas nove modalidades distintas de ensaios avaliatórios; (ii) *avaliação baseada em diretrizes de projeto e manuais de regras*; (iii) *avaliação heurística* e (iv) *métodos empíricos e outros*, conforme detalhado no Capítulo 3. Adicionalmente, foi proposta uma estratégia de avaliação adaptativa ao contexto considerado [Quei96, Quei97b].

Os resultados destes trabalhos serviram de base para o processo de mensuração do desempenho do usuário no contexto desta pesquisa, tendo sido adotada, com base nos estudos anteriores, a *observação com registro em vídeo*. De acordo com o que foi descrito na sub-seção 6.3.2, para a sondagem e estimação da satisfação do usuário no contexto desta pesquisa, foram desenvolvidos e administrados os dois instrumentos de sondagem descritos - o *DePerUSI*⁸¹ e o *OpUS*⁸².

6.4.1.8 Definição de Indicadores Objetivos e Subjetivos

A partir das técnicas de avaliação selecionadas e das metas e interesses definidos nas sub-etapas anteriores, definiu-se nesta sub-etapa os indicadores objetivos e subjetivos mais significativos para o contexto desta pesquisa. Em virtude de sua natureza híbrida, o processo avaliatório demandou um planejamento integrado de metas a serem atingidas, assim como de métricas a serem adotadas no estágio de coleta de dados, no âmbito de cada um dos enfoques considerados e, mais especificamente, daqueles envolvendo usuários de teste, tratados neste capítulo: *mensuração do desempenho e da satisfação subjetiva do usuário*.

Neste ponto, torna-se necessário um esclarecimento sobre a "aparente" *superposição* de atributos de usabilidade a serem mensurados/estimados neste enfoque híbrido, tendo em vista que o padrão selecionado para a inspeção da interface do produto considerado - o *ISO9241, Partes 11* [ISO98], *14* [ISO97a], *15* [ISO97b] e *16* [ISO99], conforme tratado no capítulo anterior, define a usabilidade de produtos em termos da efetividade⁸³, eficiência⁸⁴ e satisfação⁸⁵ com que diferentes categorias de usuários poderão alcançar as metas pré-definidas no ambiente em que o processo avaliatório será conduzido.

O modo como o *ISO9241* foi utilizado neste enfoque híbrido não implicou redundância nas bases de dados obtidas a partir dos enfoques individuais. A *Parte 11* foi empregada para nortear não apenas a *inspeção de padrões*, mas também os procedimentos individuais de *mensuração do desempenho e mensuração da satisfação*.

Por outro lado, as *Partes 14, 15 e 16*, apresentadas sob a forma de recomendações condicionais e princípios de projeto, serviram de base para a investigação de aspectos de interesse referentes, respectivamente, à comunicação usuário-produto via menus, comandos e manipulação direta.

Como indicadores quantitativos, foram adotados: (i) *tempo de conclusão da tarefa*; (ii) *número de ações incorretas*; (iii) *número de escolhas incorretas*; (iv) *número de erros repetidos* e (v) *número de consultas à ajuda*. Como indicadores qualitativos, foram consideradas: (i) *facilidade de execução da tarefa*; e (ii) *facilidade de uso do produto*.

Os Quadros 31 a 34 sumarizam a etapa de planejamento do ensaio avaliatório.

⁸¹ *Delineamento do Perfil do Usuário de Sistemas Interativos*

⁸² *Opinião do Usuário de Software*

⁸³ Precisão e completeza com que os usuários atingem metas especificadas.

⁸⁴ Recursos dispendidos com relação à precisão e completeza com que os usuários atingem metas.

⁸⁵ Ausência de desconforto e exibição de atitudes positivas no tocante ao uso do produto.

Quadro 31 – Aspectos gerais do ensaio.

ASPECTOS GERAIS DO ENSAIO	
PRODUTO	<i>MATLAB v. 5.3</i>
NATUREZA DO PRODUTO	<i>Ambiente dedicado à computação numérica e visualização de alto desempenho</i>
OBJETIVO GERAL	<i>Avaliação de aspectos do processo interativo usuário-produto</i>
OBJETIVOS ESPECÍFICOS	<ul style="list-style-type: none"> <i>i. Observação da facilidade de uso do produto;</i> <i>ii. Observação da facilidade de execução da tarefa;</i> <i>iii. Mensuração do tempo de conclusão da tarefa;</i> <i>iv. Mensuração do número de ações incorretas durante a execução da tarefa.</i> <i>v. Mensuração do número de escolhas incorretas em menus durante a execução da tarefa;</i> <i>vi. Mensuração do número de erros repetidos durante a execução da tarefa;</i> <i>vii. Mensuração do número de consultas à ajuda durante a execução da tarefa.</i>
NATUREZA DA AVALIAÇÃO	<i>Somativa objetiva/subjetiva & qualitativa/quantitativa</i>
NATUREZA DOS TESTES	<i>Campo e Laboratorial</i>
NÚMERO DE AVALIADORES	<i>02</i>
NATUREZA DAS TAREFAS AVALIADAS	<i>Mais frequentes</i>
DIMENSÃO DO UNIVERSO AMOSTRAL	<i>40 usuários de teste</i>
DURAÇÃO COMPLETA DO ENSAIO	<i>26 semanas</i> (desde o planejamento do ensaio até a apresentação dos resultados)

Quadro 32 – Aspectos específicos relativos à sondagem de campo.

ASPECTOS ESPECÍFICOS DO ENSAIO – <i>Sondagem de Campo</i>	
META	<i>Diagnóstico objetivo do processo interativo usuário-produto</i>
INTERESSE GERAL	<i>Sondagem dos perfis do universo amostral de usuários de teste e da satisfação subjetiva dos usuários no tocante ao uso do produto</i>
TÉCNICA DE AVALIAÇÃO	<i>Aplicação de Questionários (pré-teste⁸⁶ e pós-teste⁸⁷)</i>
OBJETIVOS ESPECÍFICOS	<ul style="list-style-type: none"> <i>i. Delineamento dos perfis do universo amostral de usuários de teste para sua discriminação em categorias (principiantes, intermediários e experientes);</i> <i>ii. Categorização dos usuários de teste para a elaboração das tarefas de teste do ensaio laboratorial;</i> <i>iii. Investigação da satisfação subjetiva dos usuários de teste para a confrontação com os demais fatores de avaliação considerados no experimento, i.e. desempenho do usuário como uso do produto e inspeção de conformidade do produto com padrão internacional.</i>
INDICADORES QUALITATIVOS	<i>Satisfação subjetiva</i>
DIMENSÃO DO UNIVERSO AMOSTRAL	<i>40 usuários de teste</i>
TIPOS DE RESULTADOS PREVISTOS	<ul style="list-style-type: none"> <i>i. Estudo discriminatório dos respondentes do questionário pré-teste</i> <i>ii. Bateria de tarefas de teste do ensaio laboratorial</i> <i>iii. Estudo analítico da satisfação subjetiva dos respondentes do questionário pós-teste</i>

⁸⁶ DePerUSI - Delineamento do Perfil do Usuário de Sistemas Interativos

⁸⁷ OpUS - Opinião do Usuário de Software

Quadro 33 – Síntese do planejamento da sondagem do usuário.

PLANEJAMENTO DA SONDAÇÃO DO USUÁRIO						
NATUREZA DA SONDAÇÃO	PROBLEMAS-ALVOS	OBJETIVOS DA SONDAÇÃO	TÉCNICA	FASE	INSTRUMENTO DE SONDAÇÃO	CARACTERÍSTICAS DO USUÁRIO EXPLORADAS
DE CAMPO	I	i	APLICAÇÃO DE QUESTIONÁRIOS	Pré-teste	DePerUSI ¹	Físicas
	II	ii				Relativas ao Conhecimento e à Experiência
	III	iii		Pós-teste	OpUS ²	Relativas à Tarefa e ao Trabalho
						<ul style="list-style-type: none"> ⊙ SEXO ⊙ FAIXA ETÁRIA ⊙ FORMAÇÃO ACADÊMICA ⊙ EXPERIÊNCIA COMPUTACIONAL PRÉVIA ⊙ TEMPO DE USO DE SISTEMAS COMPUTACIONAIS ⊙ FREQUÊNCIA DE USO DE SISTEMAS COMPUTACIONAIS ⊙ EXPERIÊNCIA PRÉVIA COM O PRODUTO ⊙ TEMPO DE USO DO PRODUTO ⊙ FREQUÊNCIA DE USO DO PRODUTO ⊙ EXPERIÊNCIA PRÉVIA COM PRODUTOS SIMILARES ⊙ ESPECIFICAÇÃO DE PRODUTOS SIMILARES UTILIZADOS ⊙ TEMPO DE USO DE PRODUTOS SIMILARES ⊙ NATUREZA DAS ATIVIDADES ⊙ USO CONTEXTUAL ⊙ VERSÃO UTILIZADA DO PRODUTO ⊙ FORMA DE AJUDA DO PRODUTO MAIS UTILIZADA ⊙ FACILIDADES DO PRODUTO MAIS UTILIZADAS ⊙ ASPECTOS DE USO DO PRODUTO ⊙ DOCUMENTAÇÃO <i>Online</i> e <i>Offline</i> ⊙ INTERAÇÃO GERAL COM O PRODUTO
<p>I. Com a navegação pela hierarquia de menus</p> <p>II. Com a documentação <i>Online</i> e <i>Offline</i></p> <p>III. Com a recuperação do usuário após situações de erro</p>			<p>i. Confirmação da existência de problemas</p> <p>ii. Identificação de problemas</p> <p>iii. Verificação do desempenho do usuário</p>			

¹ DePerUSI - Delimitação do Perfil do Usuário de Sistemas Interativos
² OpUS - Opinião do Usuário de Software

Quadro 34 – Aspectos específicos relativos ao teste de usabilidade em ambiente controlado.

ASPECTOS ESPECÍFICOS DO ENSAIO – <i>Teste Laboratorial</i>	
META	Diagnóstico subjetivo <i>do processo interativo usuário-produto</i>
INTERESSE GERAL	Investigação de problemas associados ao uso de comandos, à navegação pela hierarquia de menus e à eficácia das consultas à documentação.
TÉCNICA DE AVALIAÇÃO	<i>Observação Direta</i>
OBJETIVOS ESPECÍFICOS	<ul style="list-style-type: none"> i. Investigação da <i>capacidade de seleção</i> de comandos para a verificação da <i>facilidade de uso</i> pelos usuários através da janela de comandos do produto; ii. Investigação da <i>capacidade de seleção do caminho correto do menu</i> para a verificação da <i>facilidade de navegação</i> através dos menus oferecidos pelo produto; iii. Investigação da <i>facilidade de localização e compreensão da documentação</i> para a verificação da <i>facilidade de uso e eficácia das consultas à documentação</i> oferecida pelo produto.
INDICADORES QUANTITATIVOS	<ul style="list-style-type: none"> i. <i>Tempo de execução das tarefas</i> ii. <i>Número de ações incorretas</i> iii. <i>Número de escolhas incorretas no menu</i> iv. <i>Número de erros repetidos</i> v. <i>Número de consultas à ajuda</i>
INDICADORES QUALITATIVOS	<ul style="list-style-type: none"> i. <i>Facilidade de uso do produto</i> ii. <i>Facilidade de conclusão da tarefa</i>
DIMENSÃO DO UNIVERSO AMOSTRAL	40 usuários de teste
CATEGORIA DE USUÁRIOS DE TESTE	<i>principiantes (0 a 3 meses de uso), intermediários (3 meses a 1 ano de uso) e experientes (mais de 1 ano de uso)</i>
NÚMERO DE SESSÕES DE TESTE	40
DURAÇÃO DA SESSÃO DE TESTE	60 minutos
PROBLEMAS-ALVOS	<ul style="list-style-type: none"> i. <i>Com a documentação</i> ii. <i>Com o uso de comandos</i> iii. <i>Com a navegação pela hierarquia de menus</i> iv. <i>Com a recuperação do usuário após situações de erro</i>
TIPOS DE RESULTADOS PREVISTOS	<ul style="list-style-type: none"> i. <i>Confirmação da existência de problemas</i> ii. <i>Deteção de problemas</i>

6.4.2 Elaboração do Material do Ensaio

Assim como a etapa anterior, esta etapa pode ser decomposta nas seguintes sub-etapas:

6.4.2.1 Planejamento e Estruturação de Tarefas de Teste

Durante esta sub-etapa, foram ponderados (i) os *recursos humanos, físicos, materiais e temporais disponíveis* e (ii) as *metas e interesses relativos ao ensaio avaliatório*, visando o planejamento, a elaboração e estruturação das tarefas relevantes ao contexto avaliatório do produto-alvo.

É importante salientar que as tarefas de teste foram concebidas com base em contextos típicos de trabalho com o *MATLAB v.5.3.0*, visando abranger aspectos relativos à entrada, ao processamento e à saída de dados a partir do uso deste produto. Os contextos de trabalho foram analisados com base no formalismo *MAD (Méthode Analytique de Description des tâches)* [Scap89, Sebi94, Muñoz00], destinado a atividades analíticas de descrição de tarefas.

Em essência, o formalismo *MAD* é uma metodologia de análise de tarefas, numa

perspectiva de concepção de interfaces com o usuário, que integra três etapas: (i) *coleta de dados*; (ii) *descrição de tarefas*; e (iii) *validação da descrição*. Em nível da etapa de *descrição de tarefas*, as tarefas são modeladas a partir de um formalismo orientado à tarefa e adaptado aos objetivos da descrição. Vale ressaltar que o *MAD* tem sido tipicamente utilizado na etapa de análise da tarefa precedente à etapa de concepção de produtos.

No entanto, vislumbrou-se neste trabalho a possibilidade de modelar as tarefas possíveis de serem realizadas com o *MATLAB v. 5.3.0*, com o propósito de identificar "gargalos" estruturais e, por conseguinte, refinar o processo de seleção de tarefas para compor o roteiro de testes do ensaio de usabilidade planejado (mensuração do desempenho do usuário durante o uso do produto). Devido à extensão do modelo *MAD* construído para o *MATLAB v. 5.3.0* e ao seu grau de relevância para o tema central desta pesquisa, optou-se por omitir sua apresentação neste documento.

No tocante à *Tarefa 0*, também denominada *Pré-Tarefa*, foi estruturada para que se pudesse observar aspectos mais gerais do uso de sistemas computacionais e de recursos oferecidos pelo ambiente *Windows*, tais como criação de pastas, transferência de arquivos entre pastas e renomeação de arquivos.

O Quadro 35 apresenta uma síntese do planejamento das tarefas de teste.

Quadro 35 – Síntese do planejamento das tarefas de teste

PLANEJAMENTO DAS TAREFAS DE TESTE		
NATUREZA DO ENSAIO	Laboratorial	
PROBLEMAS-ALVOS	i. Com o uso de comandos ii. Com a navegação pela hierarquia de menus iii. Com a documentação iv. Com a recuperação de situações de erro	
OBJETIVO(S) DO(S) TESTE(S)	i. Confirmação da existência de problemas i. Detecção de problemas	
ESTRATÉGIA	Observação direta	
INDICADOR(ES)	Subjetivo(s)	i. facilidade de uso do produto ii. facilidade de conclusão da tarefa
	Objetivo(s)	número de ações incorretas número de escolhas erradas no menu número de erros repetidos tempo de conclusão das tarefas
NÚMERO DE TAREFAS	05	
ESPECIFICAÇÃO DAS TAREFAS	Tarefa 0 (Pré-tarefa)	Inicialização do sistema, criação de pasta de trabalho e ativação do aplicativo
	Tarefa 1	Criação de arquivos de dados
	Tarefa 2	Abertura, correção e execução de <i>script</i>
	Tarefa 3	Alteração de dados numéricos
	Tarefa 4	Configuração de representações gráficas e saída de dados numéricos e gráficos

Dentre os recursos humanos, físicos, materiais e temporais disponíveis foram discriminados nesta sub-etapa aqueles alocados para a realização de cada uma das tarefas definidas e estruturadas na sub-etapa anterior.

As Figs. 30(a) e (b) ilustram, respectivamente, concepções tridimensionais das faces anterior e posterior do *Laboratório de Interfaces Homem-Máquina (LIHM)* do *DEE/CCT/UFPB*, onde foi conduzido o ensaio de usabilidade desta pesquisa.

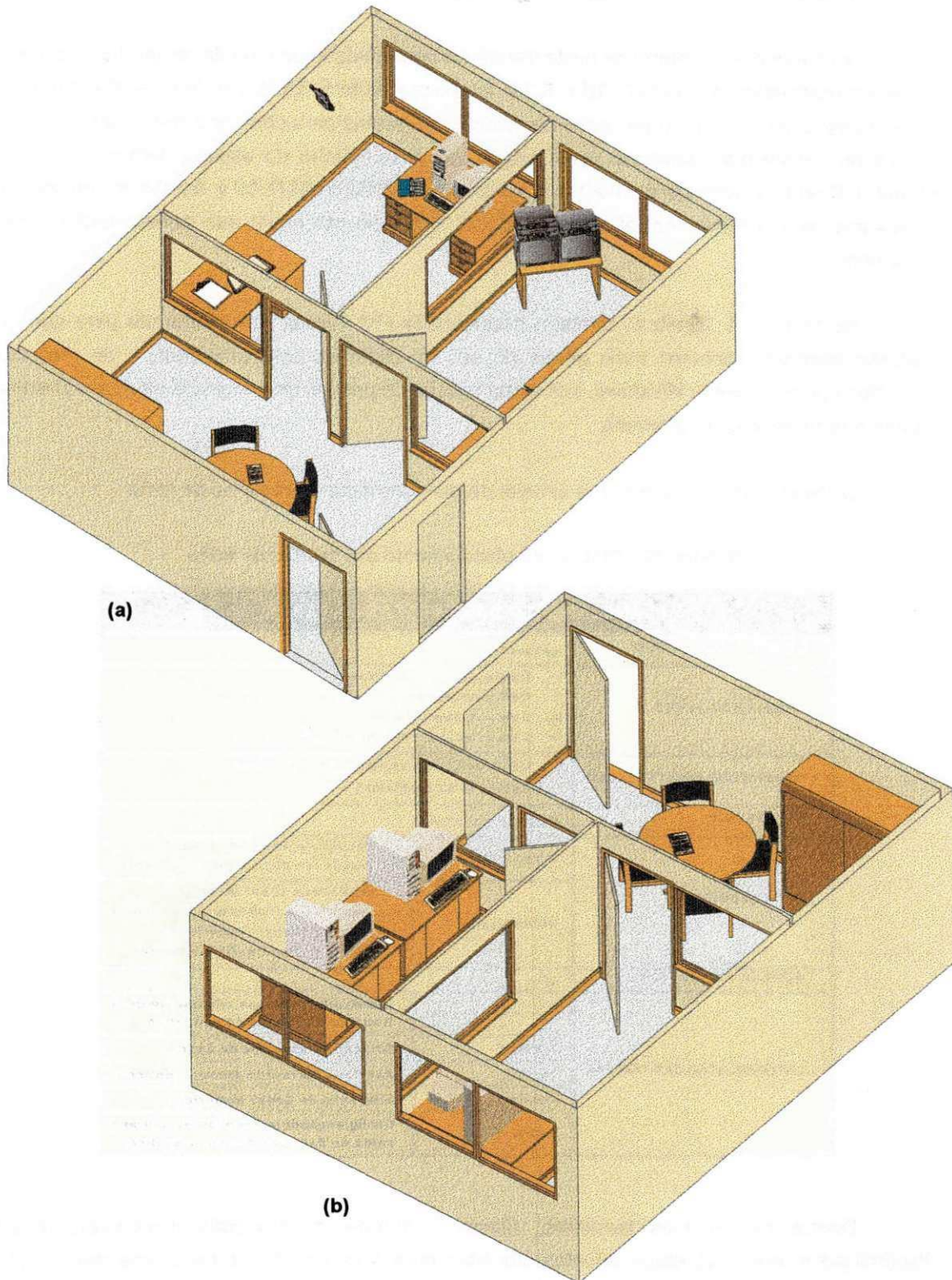


Fig. 30 – Vistas aéreas do ambiente de condução do ensaio de usabilidade (LIHM/DEE/CCT/UFPB) do ponto de vista da: (a) entrada; e (b) face posterior.

Conforme explicitado no Quadro 31, foram 02 os avaliadores participantes dos testes. Por sua vez, o Quadro 36 complementa as informações das Figs. 29 e 30, discriminando os recursos materiais (*hardware*, *software* e *outros*) utilizados durante a realização do ensaio de usabilidade, enquanto a Fig. 31 apresenta a alocação de tempo (em semanas) para a conclusão de cada fase do ensaio.

Quadro 36 – Recursos materiais utilizados no ensaio de usabilidade.

NATUREZA	ESPECIFICAÇÃO
Hardware	<ul style="list-style-type: none"> ▪ Estação de trabalho baseada em PC (2) ▪ Placa de captura de vídeo ▪ Câmaras de vídeo (2) ▪ VCR (3) ▪ Receptores de TV (2) ▪ Sistema de registro e processamento de áudio ▪ Microfones (2)
Software	<ul style="list-style-type: none"> ▪ <i>MATLAB v. 5.3.0 (The Math Works Inc.)</i> ▪ <i>Microsoft Windows95/ Office2000</i>
Outros	<ul style="list-style-type: none"> ▪ Cronômetros (02) ▪ Ficha de Cadastro do Participante ▪ Instrumentos de Sondagem (<i>DePerUSI</i> e <i>OpUS</i>) ▪ Roteiros de Tarefas (versões <i>usuário de teste</i> e <i>avaliador</i>) ▪ Fichas de Registros de Eventos ▪ Manuais do Produto Avaliado ▪ Fitas de vídeo

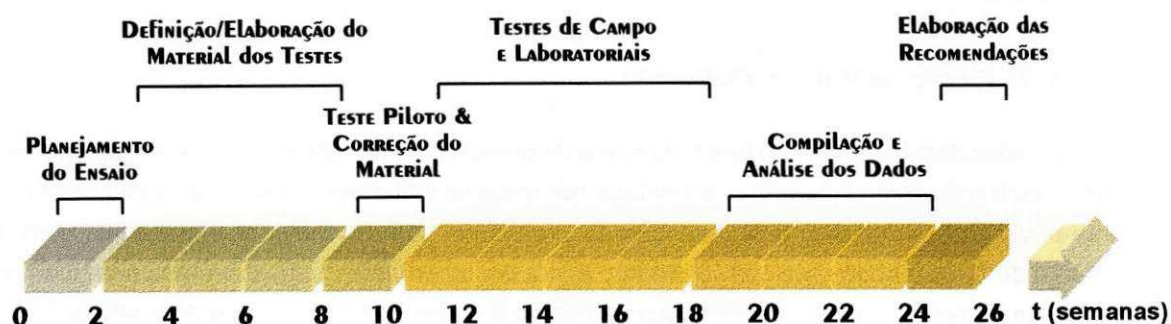


Fig. 31 – Etapas do ensaio de usabilidade.

6.4.2.3 Elaboração da Ficha Cadastral do Participante e do Documento de Aceitação das Condições de Teste

Para fins de recrutamento em experimentos futuros do *Laboratório de Interfaces Homem-Máquina* (LIHM) do DEE/UFPB, a necessidade de estruturação de um cadastro de usuários de teste potenciais, mas sobretudo para um engajamento formal dos participantes do ensaio avaliatório em questão, implicou a redação de um documento descrevendo todas as condições de teste às quais os participantes são, em geral, submetidos. Esta foi a atividade desenvolvida nesta sub-etapa, sendo a ficha cadastral apresentada no Anexo I.

Como mencionado anteriormente, o procedimento metodológico concebido por Queiroz e Turnell [Quei98a, Quei98b] contribuiu com o material necessário ao recrutamento de participantes (*Ficha Cadastral do Participante e Termo de Aceitação da Condições de Teste*), aprimorados para o contexto desta pesquisa.

6.4.2.4 Elaboração do Material Necessário à Condução do Processo Avaliatório

Destinou-se à preparação de todo o material a ser utilizado por avaliadores e usuários de teste durante o processo avaliatório do produto, assim discriminado: (i) questionários pré-teste (Anexo G) e pós-teste (Anexo H); (ii) roteiros de tarefas para o usuário de teste e para o avaliador (Anexos J e K, respectivamente); e (iii) fichas de registro de eventos (Anexo L), destinadas à anotação dos indicadores objetivos pré-definidos, assim como de comentários relativos aos indicadores subjetivos.

6.4.2.5 Discussão das Abordagens de Avaliação da Usabilidade do Produto

Este estágio da preparação do material do ensaio foi caracterizado por discussões sobre diversos detalhes inerentes ao processo avaliatório (e.g., papel e atitudes de cada membro da equipe de avaliação durante a condução do ensaio, uso adequado do material de teste, postura alternativa face à ocorrência de problemas).

Tal estratégia de ação tornou-se necessária para a avaliação da usabilidade tanto segundo os enfoques de mensuração do desempenho e da satisfação subjetiva do usuário, que exigiram uma abordagem centrada na interação usuário-produto, quanto de acordo com o enfoque da inspeção de padrões, a qual demandou a adoção de uma abordagem centrada na opinião de especialistas.

6.4.2.6 Validação do Material Elaborado

As atividades desenvolvidas ao longo da etapa de preparação do material do ensaio culminaram com a realização desta sub-etapa - a validação do material elaborado, a partir da condução de um teste-piloto, cujo propósito foi a detecção de problemas com os métodos planejados, com o material de teste elaborado e com o produto e sua documentação. Esta sub-etapa será melhor explorada no próximo capítulo, com a apresentação dos resultados do ensaio de usabilidade.

6.4.3 Condução do Ensaio e Coleta de Dados

Esta etapa, que integrou a sondagem do perfil dos participantes aos testes laboratoriais e à sondagem da opinião dos usuários sobre o produto testado, caracterizou-se pela aplicação das técnicas de avaliação da usabilidade anteriormente apresentadas, segundo as abordagens centradas na interação usuário-produto, visando a coleta dos indicadores objetivos e subjetivos que viessem respaldar a verificação da veracidade das hipóteses anteriormente formuladas.

Duas facetas caracterizaram esta etapa: a do *usuário de teste* e a do *avaliador*. Para os usuários de teste, consistiu do preenchimento do *DePerUSI*, seguido da leitura dos roteiros de tarefas^{**} (Anexo I) e da execução das seqüências de atividades pertinentes a cada um deles, encerrando com o preenchimento do *OpUS*, imediatamente após a execução da bateria de tarefas de teste.

Para o observador, esta sub-etapa correspondeu ao: (i) recebimento dos instrumentos de sondagem do perfil dos participantes; (ii) acompanhamento observatório da execução dos roteiros de tarefas pelos usuários de teste (a partir do uso dos roteiros apresentados nos Anexo K); (iii) registro dos indicadores objetivos pré-definidos, assim como de comentários relativos aos indicadores subjetivos, em fichas de registro de eventos (Anexo L); e (iv) recebimento dos instrumentos de sondagem da opinião dos usuários de teste sobre o produto testado.

6.4.4 Tabulação e Análise dos Dados

Por conveniência, a etapa de tabulação e análise dos dados referentes ao ensaio de usabilidade aparece, nesta seção, integrada àquela relativa à inspeção de conformidade do produto avaliado às partes 14, 15 e 16 do padrão *ISO 9241*.

Esta etapa, cujos resultados serão apresentados e discutidos no próximo capítulo, foi dividida em quatro sub-etapas, assim discriminadas:

6.4.4.1 Triagem Preliminar dos Dados Coletados

Contagens de tempos de execução de tarefas, de erros cometidos e repetidos, de consultas a diferentes mecanismos de ajuda e outras medidas de desempenho, opiniões, comentários verbais e escritos, registros de escalas avaliatórias semântico-numéricas e demais medidas de satisfação (assim como também anotações de problemas detectados a partir de inspeções de padrões) constituem categorias de dados que refletem, em diferentes graus e perspectivas, a usabilidade de um produto.

A partir de dados dessas naturezas, coletados na etapa anterior, foi executada uma triagem preliminar, visando a detecção de problemas colaterais, não evidenciados através de uma leitura direta dos dados (e.g., discrepâncias entre as reações do usuário durante o processo observatório e suas respostas aos itens do questionário aplicado [Quei98a, Quei98b]).

^{**} Previamente elaborados de modo a serem representativos dos diferentes interesses da investigação.

6.4.4.2 Triangulação dos Dados Coletados

Após o processo de triagem preliminar dos dados, efetuou-se a confrontação das três categorias de dados coletados (dados relativos ao desempenho do usuário, à satisfação do usuário e a conformação do produto aos padrões considerados), visando a detecção de problemas colaterais adicionais não evidenciados durante a triagem isolada dos dados obtidos a partir de cada enfoque considerado (e.g., discrepâncias entre o desempenho do usuário durante a execução das tarefas de teste e suas respostas a itens do questionário relativos a problemas encontrados na finalização das referidas tarefas).

6.4.4.3 Tabulação e Síntese dos Dados Coletados

Na verdade, o processo de tabulação dos dados coletados foi iniciado desde a primeira sub-etapa da etapa de tabulação e análise dos dados. Os resultados da triagem preliminar e da triangulação serviram de base para a conclusão do processo de tabulação dos dados, facilitando a tarefa de síntese representativa que acompanha qualquer processo de (i) tabulação de valores provenientes de coletas objetivas de dados (contagens de tempos, de erros e de pontos em escalas numéricas); (ii) listagem de dados associados a coletas subjetivas (conceitos atribuídos em escalas semânticas e problemas identificados a partir de observações).

6.4.4.4 Organização dos Problemas Listados

Nesta sub-etapa, os dados tabulados e/ou listados de forma sintética possibilitaram a organização dos problemas evidenciados a partir da condução dos processos avaliatórios, segundo o grau de severidade e de abrangência e da frequência de ocorrência, conforme será discutido no próximo capítulo.

6.4.5 Apresentação dos Resultados

Sendo a etapa final do processo avaliatório, a apresentação dos resultados foi direcionada para (i) a definição do modo de divulgação dos resultados obtidos com a aplicação das técnicas adotadas nos enfoques mencionados no capítulo anterior e neste capítulo, visando a comprovação da veracidade das hipóteses formuladas; (ii) a priorização dos dados, apresentados conforme a relevância face à comprovação das hipóteses; e (iii) a elaboração deste documento, em especial os dois próximos capítulos, que contêm a apresentação e discussão dos resultados da pesquisa (Capítulo 7) e as conclusões advindas da análise dos referidos resultados (Capítulo 8).

7

APRESENTAÇÃO E DISCUSSÃO dos Resultados

"Statistical inferences cannot be made without random samples from the populations, and random sampling from the populations of interest to the experimenter is likely to be impossible. Thus [...] the main burden of generalizing from experiments always has been, and continues to be, carried by nonstatistical rather than statistical logic."

[Eugene S. Edgington, *Randomization Tests* (1995, 3rd ed.)]

7	Apresentação e Discussão dos Resultados	211
7.1	Tipos de Dados Obtidos.....	213
7.2	Resultados da Inspeção de Conformidade do Produto com o Padrão Internacional ISO 9241	214
7.2.1	Problemas Identificados a partir da Inspeção da Conformidade do Produto à ISO 9241-14	216
7.2.2	Problemas Identificados a partir da Inspeção da Conformidade do Produto à ISO 9241-15	221
7.2.3	Problemas Identificados a partir da Inspeção da Conformidade do Produto à ISO 9241-16	228
7.2.3	Taxas de Adoção e Parecer sobre o Produto com base no Processo de Inspeção.....	236
7.3	Resultados da Mensuração do Desempenho	239
7.3.1	Análise Preliminar dos Indicadores Quantitativos	239
7.3.2	Análise Estatística dos Indicadores Quantitativos	249
7.3.3	Problemas Identificados a partir da Mensuração do Desempenho do Usuário.....	258
7.3.4	Parecer sobre o Produto com base na Mensuração do Desempenho do Usuário.....	268
7.4	Resultados da Sondagem da Opinião do Usuário	271
7.4.1	Resultados da Análise do DePerUSI	271
7.4.2	Resultados da Análise do OpUS	284
7.4.2.1	Análise Preliminar dos Escores Obtidos.....	289
7.4.2.2	Análise Estatística dos Indicadores Quantitativos	291
7.4.2.3	Problemas Identificados a partir da Sondagem da Satisfação do Usuário	294
7.5	Síntese dos Resultados Apresentados.....	298

Este capítulo apresenta os resultados obtidos a partir da aplicação dos procedimentos metodológicos descritos nos dois capítulos precedentes, discutindo-os isolada e conjuntamente. A seção introdutória – **Tipos de Dados Obtidos** – retoma sinopticamente aspectos já comentados sobre a natureza dos dados coletados a partir dos diferentes enfoques

avaliatórios adotados no âmbito desta pesquisa, estabelecendo um elo com os capítulos anteriores.

A seção 7.2 (**Resultados da Inspeção de Conformidade do Produto com o Padrão Internacional ISO 9241**) contém as listas de inspeção da conformidade do produto-alvo às *Partes 14, 15 e 16* do padrão *ISO 9241*, assim como a discussão dos problemas de usabilidade verificados no processo de inspeção e as recomendações relativas à otimização do produto face às falhas constatadas a partir deste enfoque avaliatório.

Por sua vez, a seção 7.3, **Resultados da Avaliação do Desempenho**, apresenta os dados coletados a partir dos testes de usabilidade conduzidos do *Laboratório de Interfaces Homem-Máquina* do *DEE/CCT/UFPB*, assim como o processamento estatístico desses dados, uma discussão dos problemas registrados no processo de mensuração do desempenho do usuário durante o uso do produto e as recomendações relativas à otimização do produto a partir deste enfoque avaliatório.

Na seção 7.4, **Resultados da Sondagem da Satisfação Subjetiva**, são delineados os perfis das diferentes categorias de usuários consideradas nesta pesquisa e apresentados os resultados referentes às opiniões dos usuários sobre o produto-alvo. Sobre estas informações, são discutidos aspectos referentes à satisfação subjetiva expressa através dos instrumentos de sondagem, respaldados pelo processo de observação e pelas análises retrospectivas dos registros em vídeo das sessões de teste.

Finalmente, a última seção deste capítulo (**Síntese dos Resultados Apresentados**) trata da confrontação dos resultados obtidos a partir dos diferentes enfoques avaliatórios adotados na pesquisa, resgatando as hipóteses expressas no capítulo inicial e discutindo as similaridades e as discrepâncias resultantes dos processos de (i) inspeção da conformidade do produto às *Partes 14, 15 e 16* do padrão *ISO 9241*, (ii) mensuração do desempenho do usuário e (iii) sondagem da satisfação subjetiva do usuário.

7.1 Tipos de Dados Obtidos

Ao longo da condução dos diferentes enfoques avaliatórios adotados, foram coletados vários tipos de dados, sintetizados no Quadro 37.

Quadro 37 – Tipos de Dados Coletados.

ENFOQUE AVALIATÓRIO	DADO	NATUREZA
Inspeção de Conformidade ao ISO 9241	✦ <i>verificação de recomendações individuais</i>	Qualitativa
Sondagem da Satisfação	✦ <i>características do usuário</i>	Qualitativa
	✦ <i>opiniões do usuário sobre o produto</i>	
Mensuração do Desempenho	✦ <i>tempo de execução da tarefa</i>	Quantitativa
	✦ <i>número de ações incorretas</i>	
	✦ <i>número de opções incorretas</i>	
	✦ <i>número de erros repetidos</i>	
	✦ <i>número de consultas à ajuda</i>	

Conforme descrito no Capítulo 5, o processo de inspeção da conformidade do produto-alvo às partes 14, 15 e 16 do padrão internacional ISO 9241 gerou diretamente três séries de dados qualitativos, referentes à verificação das recomendações individuais contidas em cada uma das partes utilizadas. Por sua vez, a análise desses dados gerou duas séries de dados, uma de natureza quantitativa e outra de natureza qualitativa, a saber: (i) *taxas de adoção das recomendações* e (ii) *pareceres parciais e global sobre o produto*.

A sondagem do perfil dos participantes do ensaio de usabilidade, descrita no capítulo anterior, propiciou a coleta de séries de dados qualitativos concernentes a características (i) *físicas*, (ii) *associadas ao conhecimento e à experiência* e (iii) *associadas à tarefa e ao trabalho* do usuário de teste. Por outro lado, a sondagem da opinião dos usuários de teste sobre o produto avaliado produziu séries de dados qualitativos referentes (i) *ao uso e navegação*, (ii) *à documentação online e offline* e (iii) *a impressões pessoais*.

Por fim, o processo de mensuração do desempenho, também descrito no capítulo anterior, proporcionou a aquisição de séries de dados de natureza quantitativa (vide Quadro 37), relativos aos indicadores objetivos definidos na etapa de planejamento do ensaio de usabilidade (sub-seção 6.4.1.8).

Adicionalmente, após a conclusão da etapa de coleta de dados, foram conduzidas entrevistas informais com participantes do ensaio de usabilidade. Tais entrevistas não proporcionaram *per se* a aquisição de novos dados, motivo pelo qual não figuram no Quadro 37. Entretanto, serviram para esclarecer alguns aspectos dos enfoques avaliatórios considerados, dentre os quais merecem destaque (i) *comentários feitos pelos participantes durante as sessões de teste*, (ii) *ações dos participantes registradas em vídeo e posteriormente analisadas* e (iii) *questões referentes à inspeção de conformidade do produto avaliado ao padrão ISO 9241*.

Cada uma das três seções seguintes apresenta e discute os dados obtidos a partir de cada um dos três enfoques considerados nesta pesquisa.

7.2 Resultados da Inspeção de Conformidade do Produto com o Padrão Internacional ISO 9241

Conforme comentado na sub-seção 5.5.3, o Anexo A de cada uma das partes do padrão ISO 9241 apresenta ao final uma lista de inspeção (*checklist*) estruturada em formato tabular (Table A.1) que sintetiza as recomendações contidas em cada parte do padrão, servindo de auxílio tanto para projetistas, no que diz respeito à aplicabilidade das recomendações condicionais daquela parte a contextos específicos de projeto, quanto para avaliadores de sistemas finalizados, no que concerne à verificação da adoção das recomendações pelos projetistas.

Em cada uma das listas de inspeção, há campos em branco, que devem ser preenchidos conforme a evolução do processo de inspeção. Há também campos não preenchíveis pelo projetista/avaliador, destacados dos demais pelo preenchimento em cor cinza, o que indica que não devem ser assinalados, por não se aplicarem ao processo de inspeção ou por não serem pertinentes ao contexto individual da declaração à qual estão associados.

As versões em português das listas de inspeção utilizadas nesta pesquisa, dada sua extensão⁸⁹, constituem o Anexo E, mencionados no Capítulo 5. Para a discussão do processo de inspeção da conformidade do *MATLAB* v. 5.3.0 às partes 14, 15 e 16 do ISO 9241, destacou-se no Quadro 38, extraído das listas acima mencionadas, apenas os itens para os quais o produto falhou no processo de inspeção.

Os problemas de usabilidade do produto referentes às partes consideradas do ISO 9241 aparecerão integrados no Quadro 40, apresentado na última sub-seção desta seção, a fim de dar uma visão sinóptica dos problemas detectados a partir do processo de inspeção de conformidade. A verificação da determinação das *taxas de adoção* (*adherence ratings*) das recomendações das partes 14, 15 e 16 do ISO 9241 pelo produto avaliado será efetuada a partir de consultas às listas de inspeção contidas nos anexos acima referidos, conforme se descreverá na sub-seção 7.2.4.

Também serão discriminadas na sub-seção 7.2.4, através da Tabela 1, as *taxas de adoção* (*adherence ratings*) relativas a cada uma das partes consideradas do ISO 9241, determinadas conforme descrito no procedimento metodológico apresentado no Capítulo 5. No tocante às falhas do produto-alvo associadas às recomendações das partes 14, 15 e 16 do ISO 9241 destacadas no Quadro 38, estas serão discutidas nas três próximas sub-seções e, quando se fizer pertinente, serão evidenciadas por representações gráficas das telas a partir das quais os problemas foram explicitados, com comentários que sumariam a discussão conduzida ao longo dos parágrafos de cada sub-seção. Na seqüência ordenada de comentários que acompanham cada representação gráfica dos problemas de usabilidade identificados, o quadro de abertura de de cada seqüência de comentários aparece destacado (fundo azul claro) dos demais (fundo amarelo claro), enquanto o fluxo da leitura da seqüência de comentários é indicado por setas.

⁸⁹ Na versão inicial, em formato A4, as listas resultantes da inspeção de conformidade do *MATLAB* v. 5.3.0 às partes 14 (11 pp), 15 (6 pp) e 16 (5 pp) do ISO 9241 ocupavam 22 páginas, razão pela qual se optou pela versão mais compacta apresentada no Anexo F.

Quadro 38 – Compilação das recomendações das Partes 14, 15 e 16 do ISO 9241 não adotadas pelo MATLAB v. 5.3.0.

PARTE DO ISO 9241	RECOMENDAÇÕES NÃO ADOTADAS	MÉTODO USADO NA VERIFICAÇÃO DA ADOÇÃO	
14	6.1.3 Técnicas gráficas	✦ Avaliação Analítica	
	7.2.6 Designadores numéricos		
	7.4.2 Opções em linhas – sub-ítem a)	✦ Observação	
	7.4.3 Grupos de opções		
15	5.1.3 Orientação ao usuário	✦ Avaliação Analítica	
	5.1.5 Comprimento da palavra de comando		
	5.2.1 Gerais	✦ Observação	
	Regras de abreviação		
	5.2.2	a) Regra simples	✦ Avaliação Analítica
		b) Truncamento	
	5.3.1 Gerais	✦ Observação & Avaliação Analítica	
	6.4 Correção de erros		
	6.6 Erros de digitação e ortografia (<i>misspellings</i>)		
	6.8		Comandos destrutivos
			a) Reversão (<i>undo</i>)
b) Confirmação			
7.2 c) Status do processamento			
7.4 Evidenciação de erros			
16	5.4.1 Dispositivos alternativos	✦ Observação & Avaliação Analítica	
	5.4.2 Técnicas equivalentes de manipulação direta via teclado		
	6.1.2 Sequência de manipulação direta de objetos		
	6.1.5 Retorno a estados anteriores a manipulações diretas		
	6.3.1 Visualização do arrasto		
	6.3.3 Diferenciação semântica no processo de arrasto		
	6.4.2 Mecanismos de dimensionamento		
	6.4.5 Manipulação de escala		
	6.4.6 Manipulação direta do fator de escala		
	8.2.1 Rearranjo do conteúdo visualizado de uma janela de acordo com a seleção do usuário		

7.2.1 Problemas Identificados a partir da Inspeção da Conformidade do Produto à ISO 9241-14

No tocante à recomendação 6.1.3 (Técnicas gráficas) [ISO97a], o **MATLAB v. 5.3.0 FALHA**, tendo em vista que se pode considerar como *técnica gráfica* o recurso de seleção de opções da janela de ajuda via *mouse*, como ocorre, por exemplo, na janela de ajuda do *Windows Explorer*, quando configurado para apresentação das informações como em uma página da Web. Esta estratégia também é empregada por aplicativos que adotam o estilo do ambiente *Windows* (e.g., *Globalink Power Translator Pro v. 6.2*, *Print Artist Platinum 4.0*). Uma análise comparativa das Figs. 32 e 33 ilustra melhor a falha mencionada.

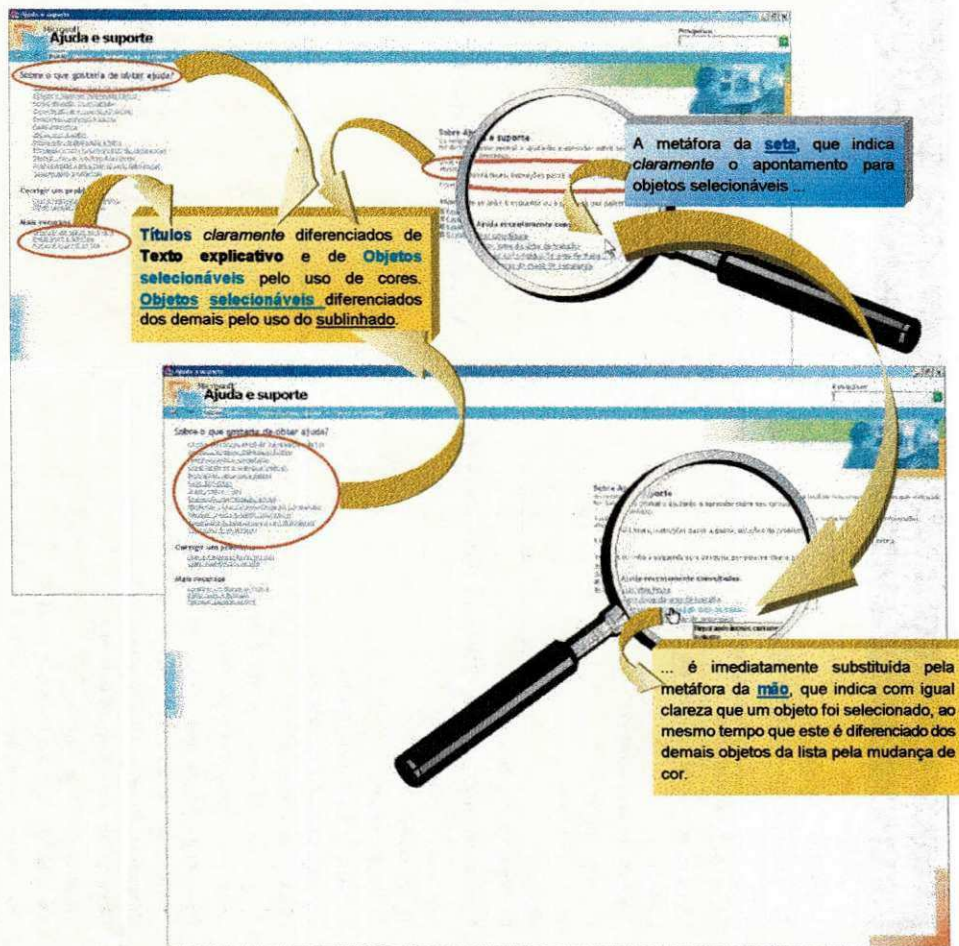


Fig. 32 – Técnica gráfica utilizada pela interface do *Windows Explorer* para a indicação das ações do usuário no processo de seleção de itens de interesse da Ajuda.

A recomendação 6.1.3 (Técnicas gráficas) da *Parte 14* do padrão *ISO 9241* contém uma nota delimitando o uso do termo *técnicas gráficas* (*graphic techniques*) ao *layout* gráfico do menu, a fim de não confundí-lo com o termo *interfaces gráficas com o usuário* (*GUI - graphical user interfaces*), que se refere usualmente a outros aspectos (e.g., dispositivos e modos de entrada e saída de dados) da interface além dos elementos gráficos. Entretanto, tal delimitação do termo transcende o realce da informação e sua discriminação pelo usuário a partir do uso de cores, tipos de linhas ou estilos de letras. O termo envolve não apenas o arranjo espacial das opções e seu agrupamento por similaridade de funcionalidades, mas também a apresentação da informação e seu acesso em nível das janelas abertas a partir da seleção dessas opções.

A Fig. 32 ilustra a janela aberta a partir da seleção da sub-opção *Tópicos da Ajuda* da opção *Ajuda* do *Windows Explorer*, o aplicativo de exploração do ambiente para o qual a versão 5.3.0 do *MATLAB* foi implementada. Tal ilustração tem apenas o propósito de evidenciar aspectos do estilo *Windows*, segundo o qual o produto avaliado foi desenvolvido. Apesar de não se estar inspecionando a conformidade com o estilo *Windows*, tal consideração se encontra implícita no processo de inspeção da conformidade do produto com o padrão internacional *ISO 9241*, visto que um dos estágios da linha evolutiva dos padrões internacionais é o estágio dos guias de estilo de plataformas industriais.

Retornando à Fig. 32, observa-se que a informação é discriminada em dois níveis mais abrangentes, a saber: (i) *objetos não selecionáveis*, que ainda podem ser subdivididos em *títulos* e *textos explicativos*; e (ii) *objetos selecionáveis*, listados como opções que podem ser acessadas conforme o interesse do usuário. O destaque e a discriminação da informação se dá tanto pelo uso do sublinhado quanto pelo uso de um grupo bem definido de cores e de tamanhos de letras (um dos atributos do estilo da letra), enquanto as alterações do *status* da seleção e dos objetos selecionáveis podem ser visualizadas, respectivamente, pela mudança do cursor de *seta* (↖) para *mão* (✎) e pela alteração da cor do objeto selecionado. As vinhetas que comentam estes aspectos na Fig. 32 foram incluídas para tornar mais evidente a clareza com que o usuário consegue diferenciar as ações possíveis de realizar sobre os diferentes níveis de informação presentes na janela de ajuda do *Windows Explorer*.

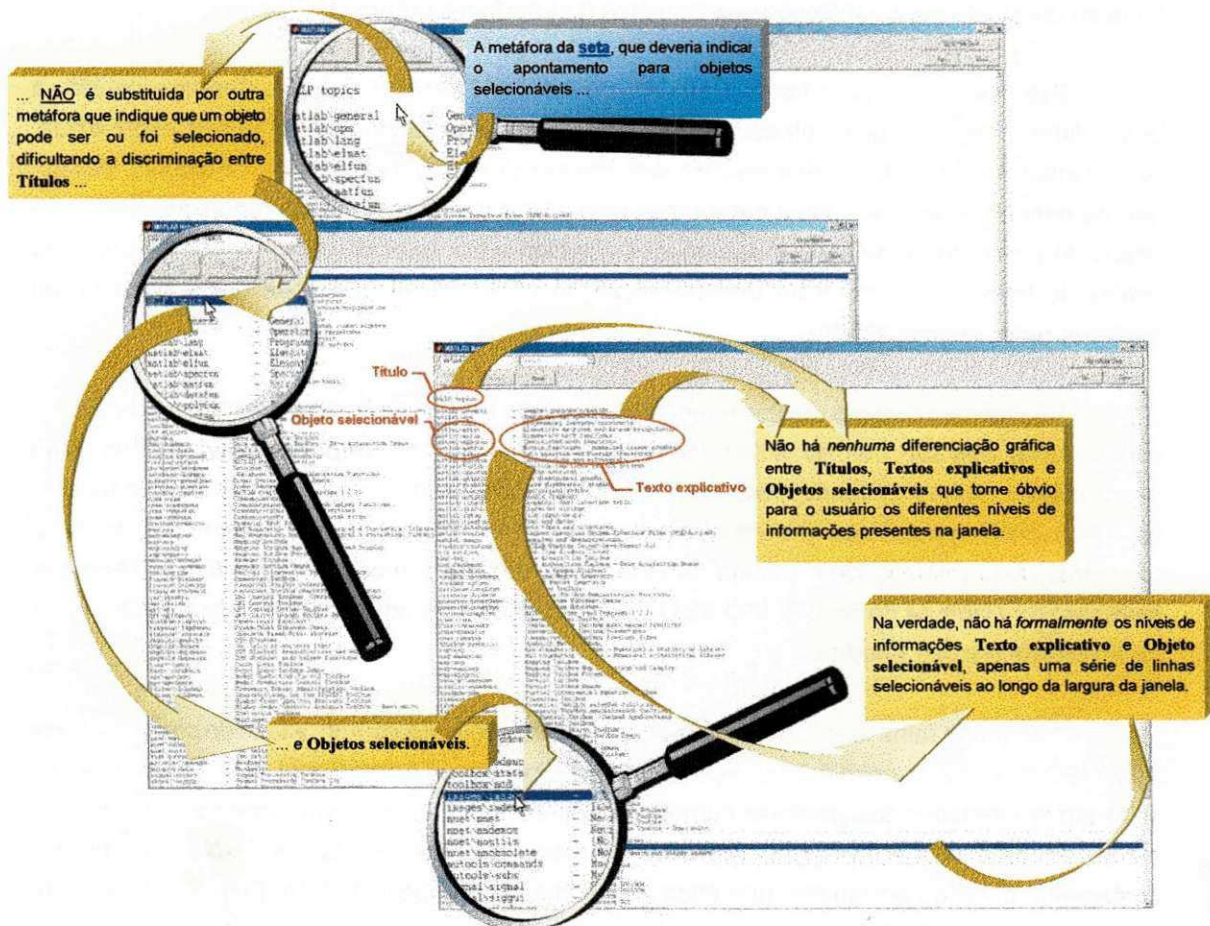


Fig. 33 – Técnica gráfica utilizada pela interface do *MATLAB* v 5.3.0 para a indicação das ações do usuário no processo de seleção de itens de interesse da Ajuda.

Observa-se, no entanto, que o processo discutido não ocorre deste modo quando o ambiente é configurado para selecionar objetos a partir do clique duplo do *mouse*. Com base neste fato, é conveniente destacar que tal configuração *ainda* permite claramente a distinção dos diferentes níveis de informação contidos numa janela no estilo *Windows*. A diferença reside apenas no fato do cursor não mudar de forma e o objeto selecionável só ser *realmente* selecionado após o clique duplo do *mouse*. Entretanto, o primeiro clique do *mouse* evidencia o objeto "pré-selecionado", invertendo a cor da letra e tarjando o objeto com uma cor de fundo diferente daquela dos demais objetos. No caso de títulos ou textos explicativos, o primeiro clique do *mouse* faz com que apenas a palavra localizada na área de apontamento seja destacada das demais e o segundo clique destaca toda a frase, como ocorre nos aplicativos de edição desenvolvidos no estilo *Windows*.

Compare-se, então, a Fig. 32 com a Fig. 33, que ilustra a janela aberta a partir da seleção da sub-opção *Help Window* da opção *Help* do *MATLAB v. 5.3.0*. Embora à primeira vista a janela *Help Window* pareça com qualquer outra janela no estilo *Windows*, há alguns aspectos que merecem ser melhor observados. O primeiro deles é o uso da área útil de apresentação dos diferentes níveis de informações de ajuda ao usuário. Percebe-se, em um primeiro momento, que praticamente *toda* a informação está concentrada do lado esquerdo da tela. Ao tentar acessá-la, vê-se, com maior clareza, que isto resulta do fato de haver apenas uma série de linhas com informações acessíveis a partir do clique duplo de um *mouse*, único fator que diferencia esta estratégia daquela adotada em interfaces textuais (o estágio evolutivo intermediário do diálogo via linha de comando para o diálogo via elementos de interfaces gráficas).

Este fato revela que a janela *Help Window* do *MATLAB v. 5.3.0* não oferece ao usuário a possibilidade de uma discriminação clara entre *objetos selecionáveis* e *textos explicativos*, simplesmente por não discriminá-los *efetivamente* em nível da implementação do recurso. De fato, não há nenhum indicativo visual entre estes dois níveis de informações nem das alterações do *status* do processo de seleção e dos objetos selecionáveis, quer em nível de uso de cores ou estilos de letras, em termos de transformação do cursor ou de destaque por sublinhado ou qualquer outro recurso similar.

Não há nem mesmo um tratamento diferenciado entre linhas selecionáveis (que conduzem a descrições das funções do *MATLAB* contidas nas listas da *Help Window*) e linhas não selecionáveis (linhas contendo *títulos* ou linhas *em branco*, usadas como separadores), dando ao usuário a falsa impressão de que um clique duplo sobre um título conduzirá a algum nível de ajuda. Em suma, esta modalidade de ajuda falha tanto no tocante à recomendação 6.1.3 (*Técnicas gráficas*) da *Parte 14* do padrão *ISO 9241 [ISO97a]* quanto na adoção de recursos gráficos que caracterizam o estilo *Windows*.

A segunda falha identificada no processo de inspeção da conformidade do produto à *Parte 14* do *ISO 9241* diz respeito à declaração 7.2.6 (*Designadores numéricos*), que recomenda que se forem empregados designadores numéricos seqüenciais, a lista deverá iniciar com o indicador 1 e não 0, como ocorre na opção *Window* do menu principal. A Fig. 34 ilustra o problema, destacando o uso do designador numérico 0 na primeira sub-opção da lista. Embora não se trate de uma falha significativa, sua correção é importante no que concerne à conformidade com o padrão *ISO 9241*.

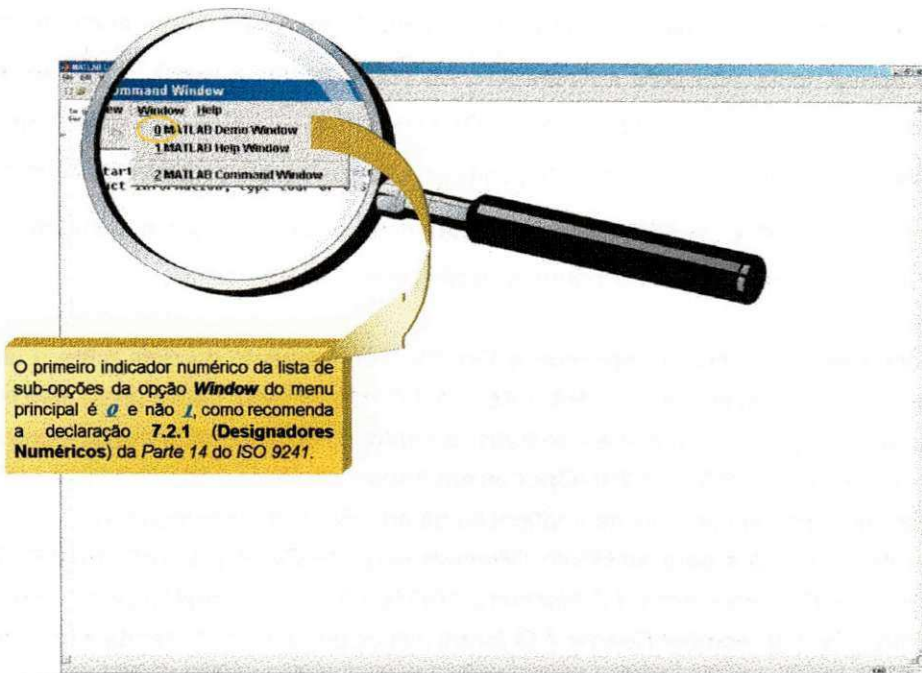


Fig. 34 – Destaque para a opção **Window** do menu principal do MATLAB e suas sub-opções.

A terceira falha identificada no produto através do processo de inspeção relativo à *Parte 14 do ISO 9241* ocorre no fechamento do ciclo do menu principal, na transição da última opção para a primeira da linha e vice-versa, conforme ilustra a Fig. 35.

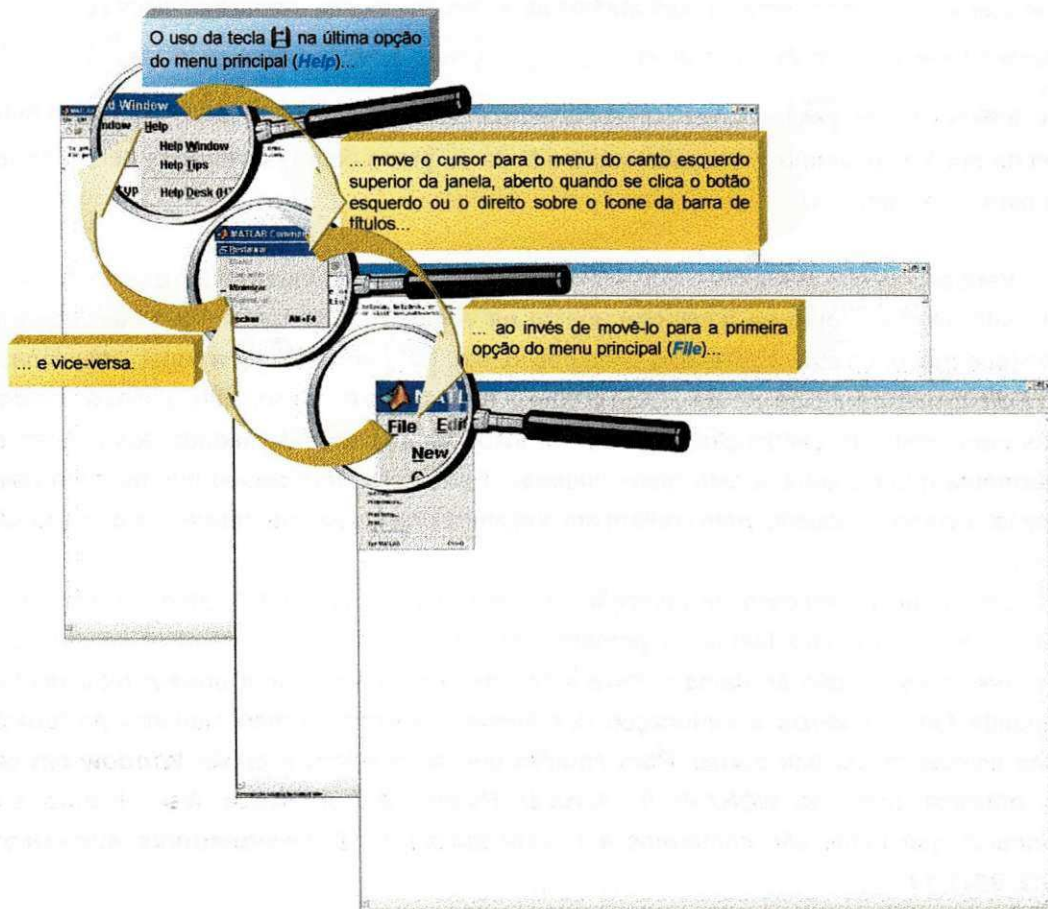




















Fig. 35 – Destaque para a falha no comportamento cíclico do menu principal do MATLAB.

Segundo a recomendação 7.4.2b) (**Opções em linhas**), quando se implementa menus cíclicos, o uso da tecla *seta para a direita* ( ou , quando a tecla  se encontra desativada) com o cursor posicionado sobre a última opção do menu de linha deverá movê-lo da última opção para a primeira opção da linha, enquanto o uso da tecla *seta para a esquerda* ( ou , quando a tecla  se encontra desativada), com o cursor sobre a primeira opção do menu de linha, deverá movê-lo da primeira opção para a última da linha.

Conforme se verifica ao observar a Fig. 35, o fechamento do ciclo entre a última e a primeira opções do menu principal do *MATLAB v. 5.3.0*. é "quebrado" com a passagem pelo menu de configuração da janela no processo de transição entre as opções extremas do menu de linha, contrariando a recomendação 7.4.2b) (**Opções em linhas**) da *Parte 14* do *ISO 9241*. Vale a pena acrescentar que, apenas para fins de verificação da adoção desta recomendação, vários outros aplicativos desenvolvidos para ambiente *Windows* (e.g. *WinZip 8.0*, *CorelDRAW 9*, *Globalink Power Translator Pro*, *Print Artist 4.0 Platinum*, *ACDSee 32 v. 2.22*, *myHouse for Windows v.6*, *Norton AntiVirus v. 5.0*, *Acrobat Reader 4.0*) foram inspecionados, tendo também apresentado a mesma falha. Adicionalmente, verificou-se também que o *MS Excel2000* também apresenta esta falha, embora os demais aplicativos do *MS Office2000*, o *MS Windows Me Explorer* e *MS Internet Explorer v. 5.50* sejam conformes à referida recomendação.

A última falha detectada a partir da inspeção de conformidade do produto à *Parte 14* do *ISO 9241* diz respeito à recomendação 7.4.3 (**Grupos de opções**). De acordo com esta recomendação, a fim de minimizar com atalhos as ações do usuário, durante a seleção por cursor, uma tecla diferente das *teclas com setas* (, ,  ou , assim como as teclas , ,  ou , quando a tecla  se encontra desativada) deverá ser usada para mover o cursor entre grupos de opções, devendo a pressão sobre a tecla ou grupo de teclas mover o cursor de um grupo para o seguinte.

Verificou-se que nenhuma tecla, além das *teclas com setas* (que movem o cursor passo-a-passo, entre opções, tanto na horizontal quanto na vertical) e das combinações de *teclas de atalhos* (que movem o cursor para uma dada opção, e.g.  +  + , a qual seleciona a opção *Edit* do menu principal do *MATLAB*) produz o movimento do cursor entre grupos de opções. Apenas para efeito de verificação, também foi inspecionada a conformidade dos aplicativos anteriormente mencionados a esta recomendação. Entretanto, verificou-se que nenhum deles oferece tal recurso ao usuário, nem contém em sua ajuda *online* qualquer referência a este tópico.

Em suma, o processo de inspeção da conformidade do produto-alvo à *Parte 14* do *ISO 9241* detectou quatro falhas no processo interativo, das quais a primeira se afigura a mais grave, com relação às demais. Para efeito de comparação, uma análise retrospectiva da segunda falha conduziu à verificação dos mesmos aplicativos inspecionados no tocante às duas últimas falhas detectadas. Para aqueles que apresentam a opção *Window* em seu menu principal (e.g., *CorelDRAW 9*, *Acrobat Reader 4.0* *myHouse for Windows v.6*) constatou-se que todos são conformes à recomendação 7.2.6 (**Designadores numéricos**) do *ISO 9241-14*.

Quanto à terceira falha, conforme relatado anteriormente, vários aplicativos inserem, no menu de linha cíclico, a passagem pelo menu de configuração da janela⁹⁰, inclusive o *MS Excel*, que constitui, sob este aspecto, uma inconsistência do *MS Office2000*. Embora tal constatação possa conduzir erroneamente à conclusão de que a recomendação **7.4.2b) (Opções em linhas)** do *ISO 9241-14* poderia merecer uma revisão, acredita-se que a recomendação é perfeitamente pertinente a qualquer contexto interativo, tendo em vista que os menus de configuração de janelas dos aplicativos não constituem uma das opções do menu principal. Se o fossem, constariam no menu de barra da tela de abertura.

Por fim, a recomendação associada à última falha identificada (**7.4.3 - Grupos de opções**), referente a transições mais rápidas entre grupos de opções de painéis de menu, merece um comentário mais detalhado. Não resta dúvida sobre a pertinência da recomendação. Se o aplicativo oferece ao usuário recursos de acesso mais rápido através de atalhos a opções e sub-opções isoladas de seu menu, se afigura lógico que também ofereça recursos similares no tocante aos grupos de opções que compõem os diferentes painéis de sub-menus.

Evidentemente, a investigação de outros aplicativos com relação a adoção desta recomendação pareceu bastante razoável, uma vez que possibilitaria uma visão mais abrangente de como esta recomendação é encarada pelos desenvolvedores, ao mesmo tempo que uma consulta aos usuários poderia esclarecer o grau de relevância deste aspecto do processo interativo para seus contextos de trabalho. Entretanto, como tal iniciativa se afigurou secundária ao escopo desta pesquisa, a investigação restringiu-se ao *MATLAB v. 5.3.0*.

7.2.2 Problemas Identificados a partir da Inspeção da Conformidade do Produto à *ISO 9241-15*

A primeira falha identificada durante a inspeção da conformidade do *MATLAB v. 5.3.0* à *ISO 9241-15* [ISO97b] é relativa à declaração **5.1.3 (Orientação ao usuário)**, segundo a qual deve-se escolher nomes de comandos que sejam consistentes com a experiência do usuário, correspondendo pois à sua linguagem operacional. Dada a abrangência da gama de interpretações que podem ser dadas a esta recomendação, a ISO acrescenta a esta recomendação uma nota referente aos produtos destinados a diferentes categorias de usuários, através da qual afirma que "pode ser importante"⁹¹ fornecer diferentes grupos de nomes de comandos como opções para diferentes categorias de usuários.

O trecho da referida nota, citado no parágrafo anterior dá, não apenas à nota, mas à recomendação **5.1.3**, um caráter de informalidade que, ao invés de delimitar, amplia a gama de interpretações possíveis para a recomendação. Na verdade, a responsabilidade pela aprovação ou reprovação referente à conformidade do produto a esta recomendação recai muito mais sobre a interpretação da declaração e da nota anexa pelo avaliador do que sobre a verificação pura e simples do produto quanto à adoção da recomendação.

⁹⁰ No caso do *myHome for Windows v. 6* (*DesignWare Inc.*), o ciclo inclui, além da passagem pelo menu de configuração da janela de abertura do aplicativo, a passagem pelo menu de configuração da janela de projeto ativa.

⁹¹ Na versão original: "*If there are multiple user groups, it may be important to provide different sets of command names for these different groups*" [ISO97b].

No processo de inspeção de conformidade do *MATLAB v. 5.3.0* a esta recomendação julgou-se pertinente, com base na observação e em análises contextuais de uso, a aplicabilidade e adoção desta recomendação pelo produto, dada a existência dos múltiplos grupos de usuários aos quais se destina, dotados de experiências e habilidades as mais diversificadas. Tendo em vista não oferecer os recursos sugeridos na nota da declaração **5.1.3**, o produto falha no tocante a esta recomendação.

A falha seguinte diz respeito à declaração **5.1.5 (Comprimento da palavra de comando)**, que recomenda que as palavras de comando não devem exceder sete caracteres. Se a inspeção tivesse se fundamentado *apenas* na recomendação propriamente dita, dir-se-ia simplesmente que o produto falha por conter em seu elenco comandos executáveis a partir de palavras que excedem sete caracteres, conforme exemplifica o Quadro 39.

Quadro 39 – Exemplos de palavras de comando do *MATLAB v. 5.3.0* que contrariam a recomendação 5.1.5 da *ISO 9241-15 [ISO97b]*.

# DE CARACTERES	EXEMPLOS
19	<i>ordergroupscallbacks, dng_simplifiedashboard</i>
18	<i>waitforbuttonpress, conditioncallbacks</i>
17	<i>powergui_loadflow, cls_sys_proptable, cls_sig_proptable, cls_mdI_proptable</i>
16	<i>selectmoveresize, psbmonophaseline, csIsortblocklist</i>
15	<i>powergui_steady, psbseriesbranch, psbttriphaseline, psbxfosaturable, chgfigproptable</i>
14	<i>demodAQ_action, demoai_channel, demoao_channel, demoai_logging, rebuffer_delay</i>
13	<i>uicontextmenu, demodAQ_intro, psbtransfosat, psbseriesload</i>
12	<i>recordsound, contourslice, reducevolume, pagesetupdlg, demodAQ_save</i>
11	<i>reducepatch, shrinkfaces, actxcontrol, uiclearmode, partialpath</i>
10	<i>reducepath, ctranspose, persistent, precedence, cholupdate</i>
9	<i>otherwise, workspace, inpolygon, Isqnonneg, fileparts</i>
8	<i>pathtool, wavwrite, whatsnew, ipermute, rtfindr</i>

Entretanto, o julgamento também foi condicionado às duas notas adicionais da declaração **5.1.5**. A primeira delas afirma que pode ser apropriado usar palavras de comando que excedam a regra dos sete caracteres, quando tais palavras forem mais naturais do que uma abreviação, exemplificando com a palavra inglesa *allocate* (alocar) e com a palavra alemã *einfügen* (inserir), ambas com 8 caracteres. A segunda nota aconselha a consulta às recomendações sobre abreviações para palavras de comando longas - seção **5.2 (Abreviações)**, composta pelas declarações **5.2.1 (Gerais)** e **5.2.2 (Regras de abreviação)**, por sua vez subdividida em **a) (Regra simples)** e **b) (Truncamento)**.

Com relação à primeira nota, o *MATLAB v. 5.3.0* “passaria” no que diz respeito a diversas palavras de comando, tais como *pathtool*, *workspace*, *persistent*, *partialpath* ou mesmo *contourslice*, tendo em vista que significam muito mais para um usuário pertencente a qualquer categoria (inclusive usuários experientes), do que palavras de comando equivalentes abreviadas ou truncadas. No entanto, ao observar novamente os exemplos do Quadro 39, verifica-se que várias das palavras de comando (e.g. *actxcontrol*, *pagesetupdlg*, *psbtransfosat*, *chgfigproptable*, *demodAQ_action*, *cls_mdI_proptable*) contrariam não somente a declaração **5.1.5**, não se respaldando na primeira nota, como também as declarações da seção **5.2 (Abreviações)**. Como

se pôde verificar, palavras como estas não só não utilizam uma regra simples de abreviação, lógica para o usuário, como também não são resultantes da aplicação de uma regra consistente de truncamento. Na verdade, muitas das palavras de comando do *MATLAB* v. 5.3.0 são um misto de truncamento de uma parte e uso integral do restante, como exemplifica *ifftshift* (um misto da truncagem da expressão *inverse fft* e da palavra *shift*) ou *cholupdate* (híbrida da truncagem do nome próprio *Cholesky* e da palavra *update*). Deste modo, o produto falha não só no tocante à recomendação 5.1.5, mas também no que concerne às recomendações 5.2.1 e 5.2.2 da ISO 9241-15 [ISO97b].

A quarta falha detectada é relativa à declaração 5.3.1 (Geral) da seção 5.3 (Teclas de funções e teclas de atalho), que recomenda para os casos de uso de teclas de funções ou de atalhos para a entrada de comandos que tal uso deverá ser óbvio para o usuário ou que as combinações de teclas deverão ser facilmente acessíveis ao usuário, mantendo sua consistência ao longo da aplicação. Esta declaração ressalva através de uma nota a importância de se considerar o uso de teclas de funções e de atalhos para comandos freqüentemente usados ou em situações nas quais a rapidez do processo de entrada de comandos é um fator relevante. Embora a versão 5.3.0 integre em sua interface três modos de interação (*linguagem de comandos, menus e manipulação direta*, conforme destacado na Fig. 36), entende-se que o usuário deve acessar TODAS as facilidades oferecidas através de qualquer um desses modos, embora também possa utilizar mecanismos de interação pertencentes aos três modos, se assim lhe for conveniente. No entanto, se for considerado apenas modo de interação via linguagem de comandos, o uso das teclas de atalho não é óbvio para o usuário, assim como não são facilmente acessíveis pelo usuário informações sobre as combinações de teclas adotadas no *MATLAB* v. 5.3.0.

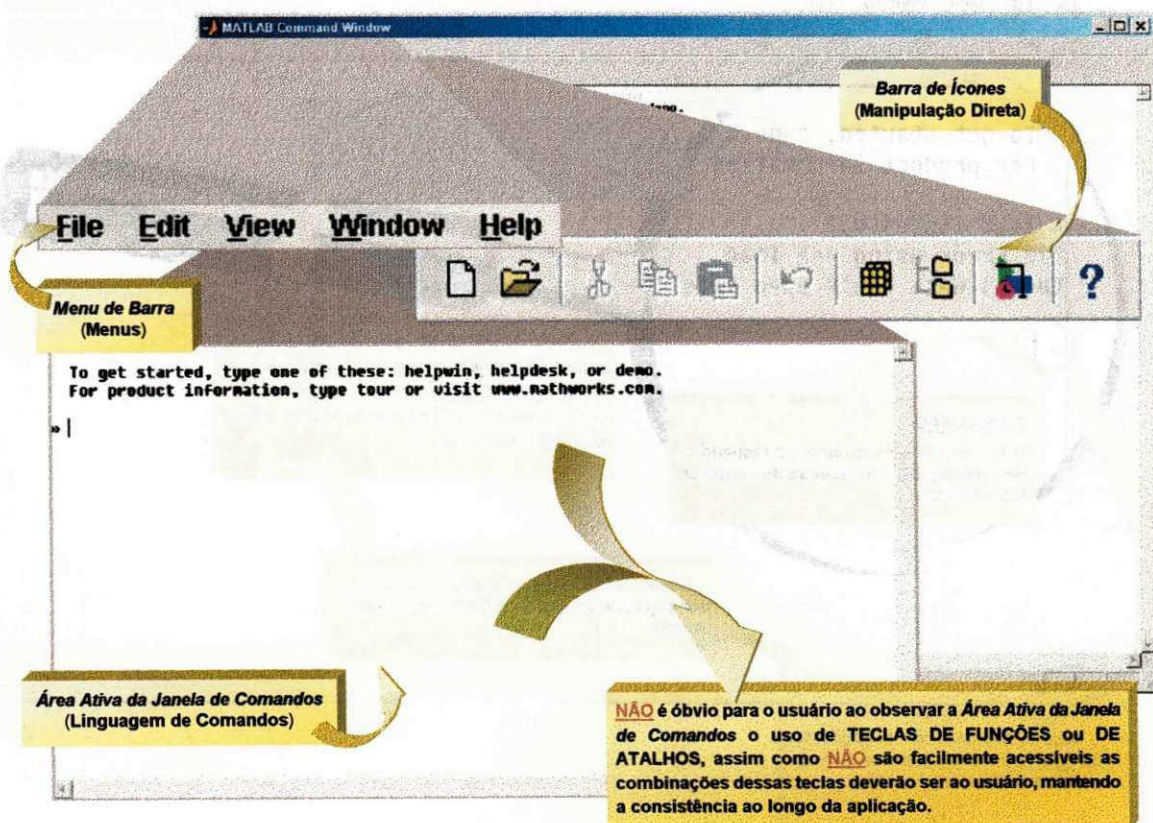










Fig. 36 – Mecanismos de interação do usuário com as facilidades oferecidas pelo *MATLAB* v. 5.3.0, com destaque para a falha referente à recomendação 5.3.1 (Geral).

Caso já tenha usado outros aplicativos desenvolvidos no estilo *Windows* com modo de interação via menus, o usuário poderá supor que a combinação das teclas  +  +  dará acesso à opção **File** do menu via teclado e, a partir daí verificar qual atalho deverá ser utilizado para acessar uma dada sub-opção (será muito menos óbvio para o usuário descobrir que a combinação  +  +  +  +  lhe permitirá criar um novo modelo).

No entanto, o tempo dispendido para identificar a combinação de interesse suprimirá a vantagem da rapidez do uso daquele atalho para acesso à função desejada. Além do mais, se assim o fizer, o usuário estará acessando recursos do modo de interação via menus como um mecanismo de compensação para a ausência de recursos equivalentes no diálogo via linguagem de comandos. A Fig. 36 ilustra a estruturação da janela de comandos (*Command Window*) do *MATLAB* v. 5.3.0, reforçando graficamente a falha do produto em relação à recomendação 5.3.1.

A quinta falha identificada através do processo de inspeção da conformidade do produto à *ISO 9241-15* refere-se à declaração 6.4 (**Correção de erros**), que recomenda, na ocorrência de erros no processo de entrada de frases de comando, a solicitação de redigitação ou edição apenas do trecho incorreto do comando e parâmetros associados. Conforme ilustra a Fig. 37, os erros cometidos pelo usuário são sucedidos por uma mensagem de erro, nem sempre suficientemente esclarecedora, após o que o aplicativo apresenta o *prompt* para a reentrada de toda a linha de comando incorreta.

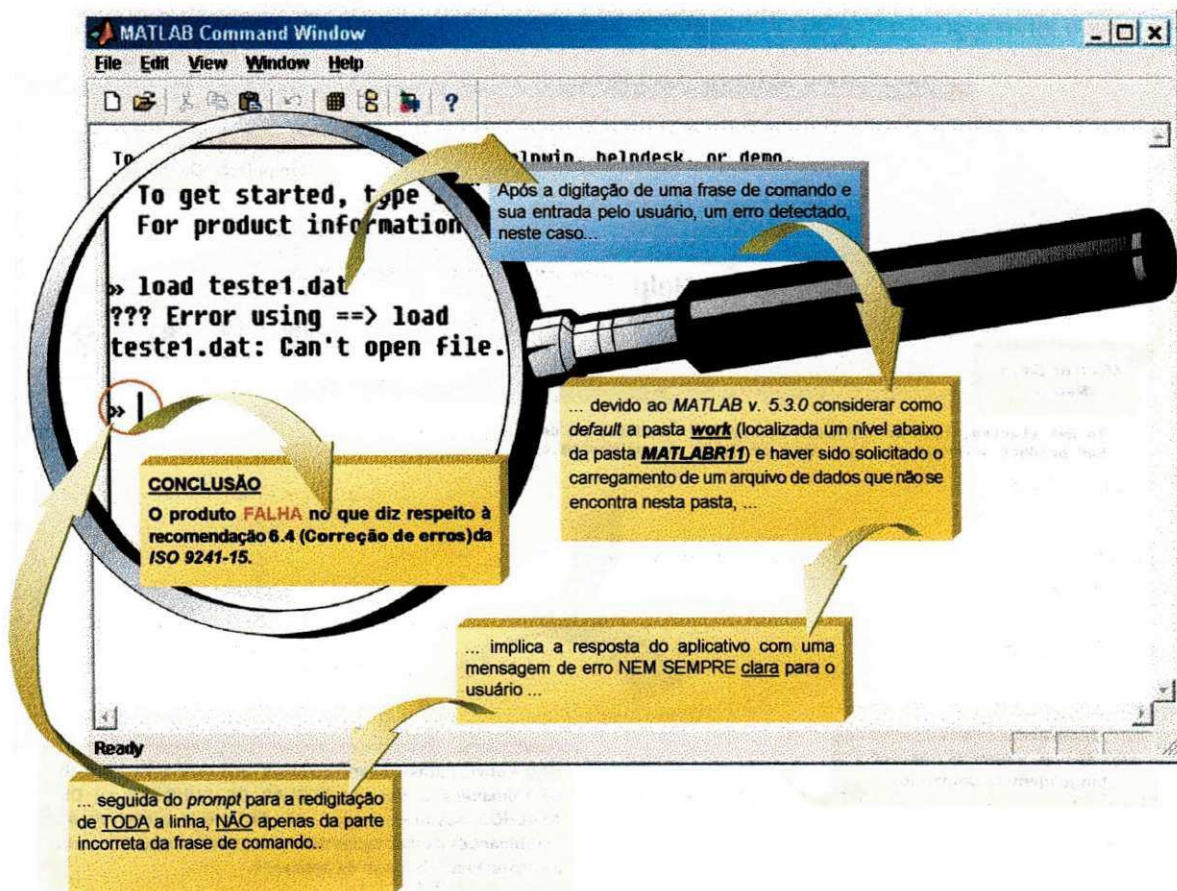


Fig. 37 – Destaque para a falha do *MATLAB* v. 5.3.0 no tocante à recomendação 6.4 (Correção de erros) da *ISO 9241-15*.

O produto inspecionado também falha com relação à recomendação 6.6 (Erros de digitação e ortografia), de acordo com a qual o sistema deverá interpretar e aceitar, se for apropriado para a tarefa e permitido pelas restrições do sistema, a entrada de comandos submetidos com erros de digitação e/ou ortografia, a menos que haja ambigüidade quanto ao comando pretendido. Contrariando esta recomendação, a interface de linha de comando do *MATLAB v. 5.3.0* **NÃO** é tolerante a erros de digitação e ortografia cometidos nas frases de comando, mesmo aos mais simples, conforme ilustra o exemplo apresentado na Fig. 38.

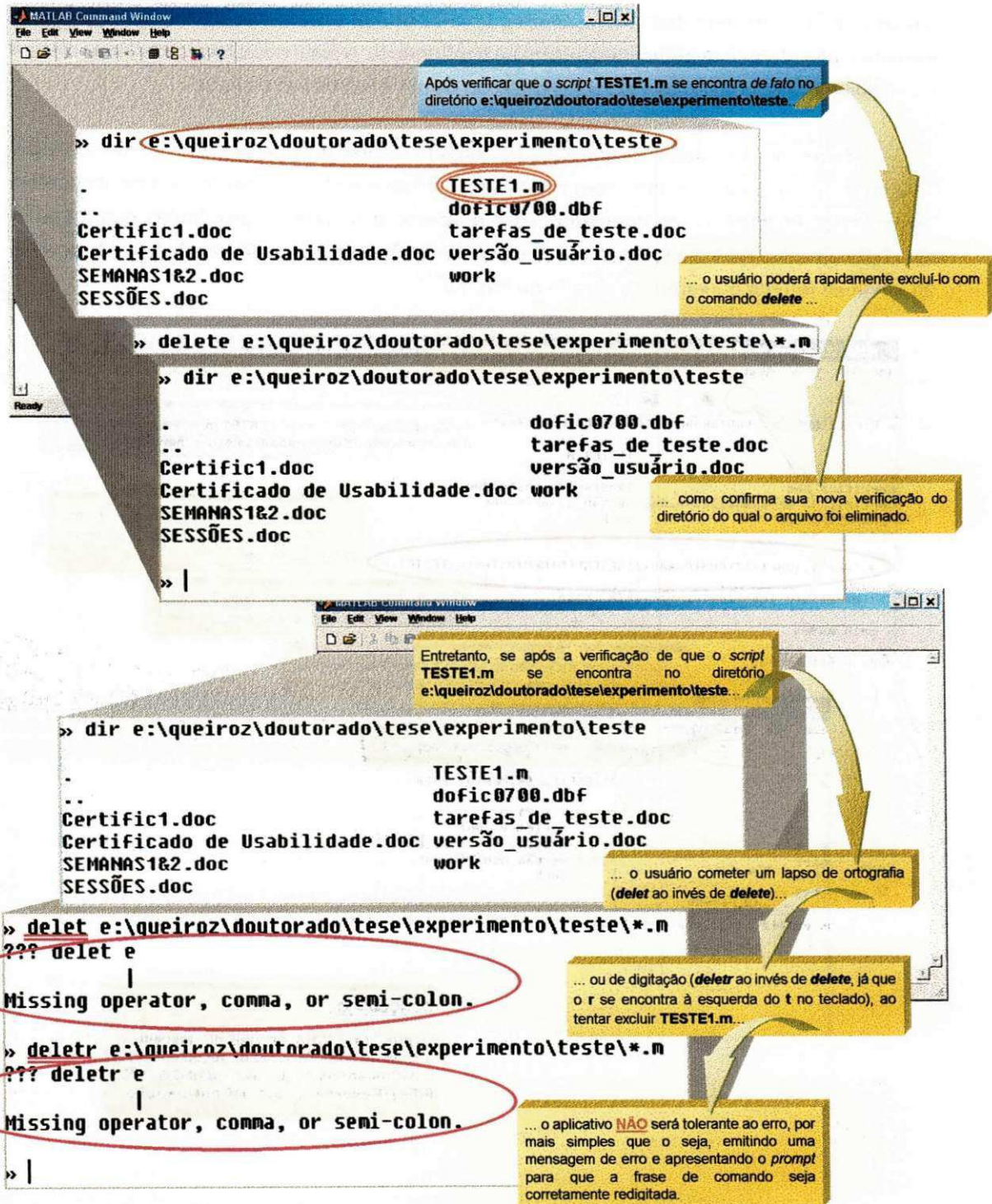


Fig. 38 – Destaque para a falha do *MATLAB v. 5.3.0* referente à correção de erros (recomendação 6.4 da ISO 9241-15)

A recomendação 6.8 (Comandos destrutivos) da ISO 9241-15 trata das estratégias de ação de um sistema face à entrada de comandos que implicam conseqüências destrutivas ou não desejadas, exemplificando o fato justamente com o apagamento de arquivos, situação focalizada no exemplo ilustrado da Fig. 38. Em situações desta natureza, a ISO recomenda que: (i) sejam disponibilizados mecanismos que permitam ao usuário reverter ao estado anterior ao último comando digitado e eliminar seus efeitos (sub-cláusula 6.8a) **Reversão**); ou (ii) seja solicitada do usuário a confirmação da execução do comando, antes de fazê-lo (sub-cláusula 6.8b) **Confirmação**). Um retorno à Fig. 38, desta feita visando a observação da ação do aplicativo à entrada de um comando destrutivo (*delete*), possibilita à primeira vista a verificação de que o aplicativo não oferece nenhum mecanismo de confirmação. A partir deste fato, poder-se-ia crer que o aplicativo oferece recursos para a reversão do último comando executado.

Entretanto, constatou-se que **NÃO** há nenhum recurso destinado à reversão de comandos destrutivos ou não intencionais, nem mesmo se o usuário tentar compensar a ausência de tais mecanismos no diálogo via comandos com o acesso a recursos equivalentes nos modo de interação via menus ou manipulação direta. Tal afirmação pode ser verificada observando-se a Fig. 39, onde se retoma o exemplo ilustrado da Fig. 38.

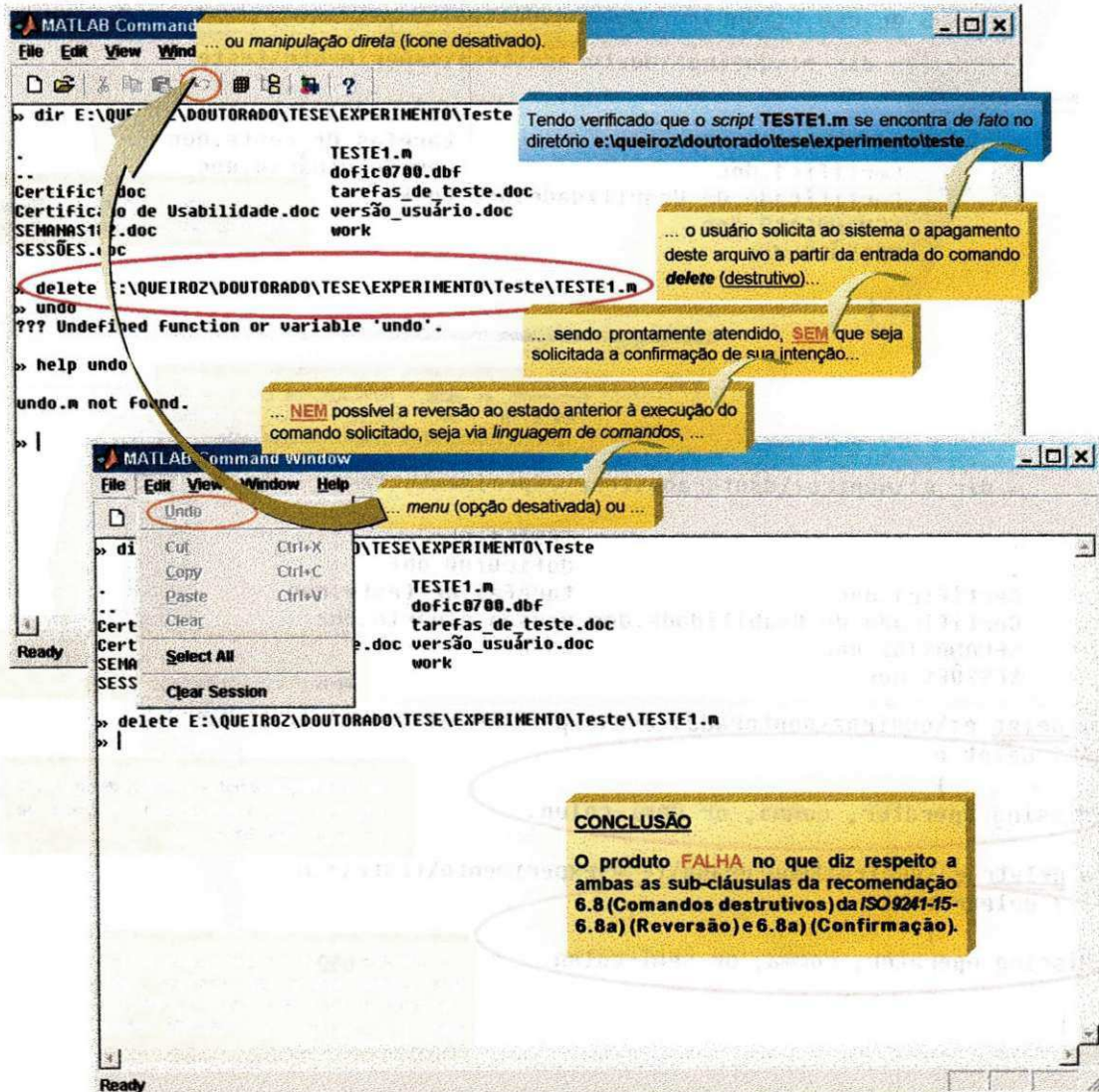


Fig. 39 – Destaque para a falha do MATLAB v. 5.3.0 referente a comandos destrutivos.

Do mesmo modo que NÃO existe uma função denominada **undo**, também não há uma função inversa ou dual da função **delete** (**undelete**). Outros exemplos de comandos destrutivos da biblioteca de comandos do *MATLAB v. 5.3.0* são **clear** e **close**, a partir dos quais o usuário pode, respectivamente, perder todo um processamento de horas ao “limpar” inadvertidamente o conteúdo do espaço de trabalho (sintaxe: **clear all**) ou fechar uma janela (sintaxe: **close**, que fecha a janela corrente de figuras) ou todo um conjunto de janelas nas quais o usuário esteja executando ações do tipo “entrada interativa via *mouse*” (sintaxe: **close all**).

Tal falha do aplicativo avaliado nos reporta a outra, apresentada aqui fora da seqüência. Trata-se da não conformidade à declaração da sub-cláusula **5.1.2d) – Pares de comandos congruentes**, que recomenda que o sistema ofereça ao usuário pares de comandos congruentes para as operações que possuam inversas ou duais (e.g., *read/write*, *sin/asin*, *do/undo*). Como se pôde observar na discussão da falha referente à recomendação **6.8 (Comandos destrutivos)**, nem todas as operações que possuem inversas ou duais foram implementadas aos pares no *MATLAB v. 5.3.0*, e.g. **delete**, **clear**), motivo pelo qual o aplicativo também falha no que diz respeito à sub-cláusula **5.1.2d)**.

No tocante à declaração **7.2 (Processamento de comandos)**, a *ISO* apresenta suas recomendações em três sub-cláusulas. A sub-cláusula **7.2b) (Retorno imediato)**, uma complementação da sub-cláusula **7.2a) (Conclusão)** no que concerne a tempos de processamento de comandos superiores a 5 segundos, recomenda a inclusão de recursos visuais que forneçam ao usuário informação sobre a ação em curso. O *MATLAB v. 5.3.0* “passa” quanto a esta recomendação, pois utiliza a metáfora visual da ampulheta para indicar “execução em andamento” nas situações em que o processamento dos comandos excede 5 segundos.

Com relação à última sub-cláusula da declaração **7.2 (7.2c) Status do processamento)**, complementar das demais, através da qual a *ISO* recomenda, caso apropriada à tarefa e permitida pelas restrições do sistema, a inclusão de mecanismos de realimentação concernente à quantidade de tempo restante para a conclusão de ações em curso. O *MATLAB v. 5.3.0* falha com relação a esta sub-cláusula por NÃO oferecer nenhuma realimentação do tempo restante para a conclusão de processos em andamento.

A última falha verificada através da inspeção de conformidade da *ISO 9241-15* concerne à declaração **7.4 (Evidenciação de erros)**, segundo a qual se a realimentação a erros for apresentada na linha de comando e se houver um erro na entrada de dados, o sistema deverá realçar a parte inaceitável do comando digitado (no contexto do comando completo ou apenas de uma porção lógica). Conforme já se pôde verificar a partir dos exemplos ilustrados anteriormente, embora as mensagens de erro oferecidas pelo *MATLAB v. 5.3.0* tentem evidenciar as partes inaceitáveis dos comandos, em diversas situações NÃO o fazem com clareza, sendo excessivamente lacônicas e de pouca ou nenhuma ajuda, sobretudo quando se considera as dificuldades normalmente enfrentadas por usuários principiantes. Se a Fig. 38 for novamente observada, ver-se-á que para duas falhas distintas de digitação – *delet*, omissão de um caractere da palavra de comando e *deletr*, digitação de um caractere incorreto na palavra de comando – o tratamento dos erros pelo aplicativo se dá através de uma única mensagem: *Missing operator, comma or semi-colon*. Esta mensagem de erro, além de lacônica, não esclarece as razões das falhas acima citadas, pois não há operadores, vírgulas ou pontos-e-vírgulas faltando na sintaxe da frase de comando.

Sumariando o processo de inspeção da conformidade do produto-alvo à *Parte 15* do *ISO 9241* possibilitou a identificação de 14 falhas no processo interativo, das quais aquela referente à recomendação **6.8 (Comandos destrutivos)** se afigura a mais grave, embora as demais, de caráter intermediário e superficial, também mereçam ser convenientemente consideradas e corrigidas. No que concerne à análise comparativa conduzida no processo de inspeção da *ISO 9241-14*, apresentada na sub-seção anterior, não foi possível por não se dispor de aplicativos dotados de interfaces com diálogo via linguagem de comandos.

7.2.3 Problemas Identificados a partir da Inspeção da Conformidade do Produto à *ISO 9241-16*

A declaração **5.4.1 (Dispositivos alternativos)** da *Parte 16* do *ISO 9241* [ISO99a] recomenda que para diálogos fundamentados em procedimentos de manipulação direta de objetos TODAS as manipulações deverão ser acessíveis ao usuário pelo menos a partir de um dispositivo de apontamento, porém se forem disponibilizados dispositivos alternativos de entrada, aquelas manipulações diretas que se adequarem melhor aos dispositivos alternativos disponibilizados também deverão ser acessíveis ao usuário a partir de tais dispositivos. Adicionalmente, a declaração **5.4.2 (Técnicas equivalentes via teclado)** recomenda para os usuários que experimentem dificuldades na operação de dispositivos de apontamento (e.g. *mouse*) que sejam oferecidas técnicas equivalentes para um teclado ou dispositivo similar que conduzam aos mesmos resultados obtidos com um dispositivo de apontamento, não necessariamente fundamentadas em manipulação direta.

A inspeção da conformidade do *MATLAB v. 5.3.0* a estas recomendações, estreitamente relacionadas entre si, conduziu à constatação de que além de não haver sido oferecido nenhum mecanismo alternativo de acesso à barra de ferramentas do produto via teclado, contrariando o que foi exposto nos dois parágrafos anteriores, não há menção a esta questão na documentação *online* ou *offline*, o que *per si* já representa uma falha adicional da documentação. A primeira constatação contraria diretamente a recomendação **5.4.2**, pois deveria ser fornecido um mecanismo de acesso à barra de ferramentas da janela principal do produto a partir do teclado. Tendo em vista que a barra de ferramentas constitui o portal de acesso aos recursos de manipulação direta oferecidos pelo produto, a ausência de mecanismo de acesso a tais recursos via teclado contraria indiretamente a recomendação **5.4.1**.

A terceira falha identificada no produto relativa a inspeção de sua conformidade à *Parte 16* do *ISO 9241* diz respeito à declaração **6.1.2 (Seqüência de manipulação direta de objetos)**, que recomenda a ordenação consistente no processo de modificação de objetos, implementada de modo que a seleção de um objeto preceda sua manipulação direta, a menos que a tarefa requiera uma ordenação alternativa dos eventos. Embora o produto seja conforme a esta recomendação na maioria das situações a ela associadas, há algumas que merecem destaque, por serem contrariadas. Novamente a falha diz respeito à janela de ajuda (*Help Window*), que também apresentou uma falha na inspeção de conformidade do *MATLAB v. 5.3.0* à *ISO 9241*. Esta janela está associada ao primeiro grupo de opções do sub-menu subordinado à opção *Help*, que integra as sub-opções *Help Window* e *Help Tips*. Em um primeiro nível, a falha é observada quando se seleciona a primeira das sub-opções mencionadas. Tendo em vista que, ao se abrir, a janela *já* exhibe o nível mais elevado da hierarquia, i.e. a estrutura de tópicos (*HELP Topics*), não se justifica

a condição *habilitada* do botão *Home*, visto que isto implica a possibilidade de seu acionamento pelo usuário e, nestas circunstâncias, tal acionamento se mostra confuso e desnecessário, não produzindo efetivamente nenhum resultado, tendo em vista que o usuário *já* está no nível de informação para o qual o botão, ao ser acionado, o conduziria (vide Fig. 40).

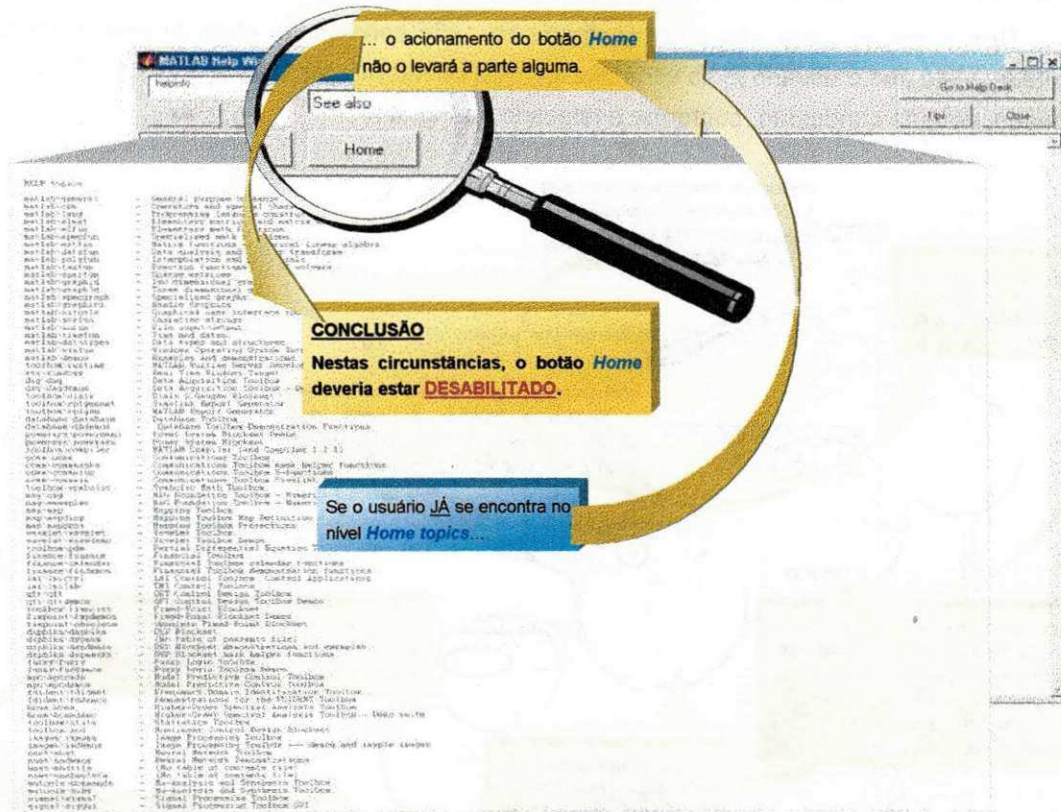


Fig. 40 – Destaque para a falha no comportamento do botão *Home*.

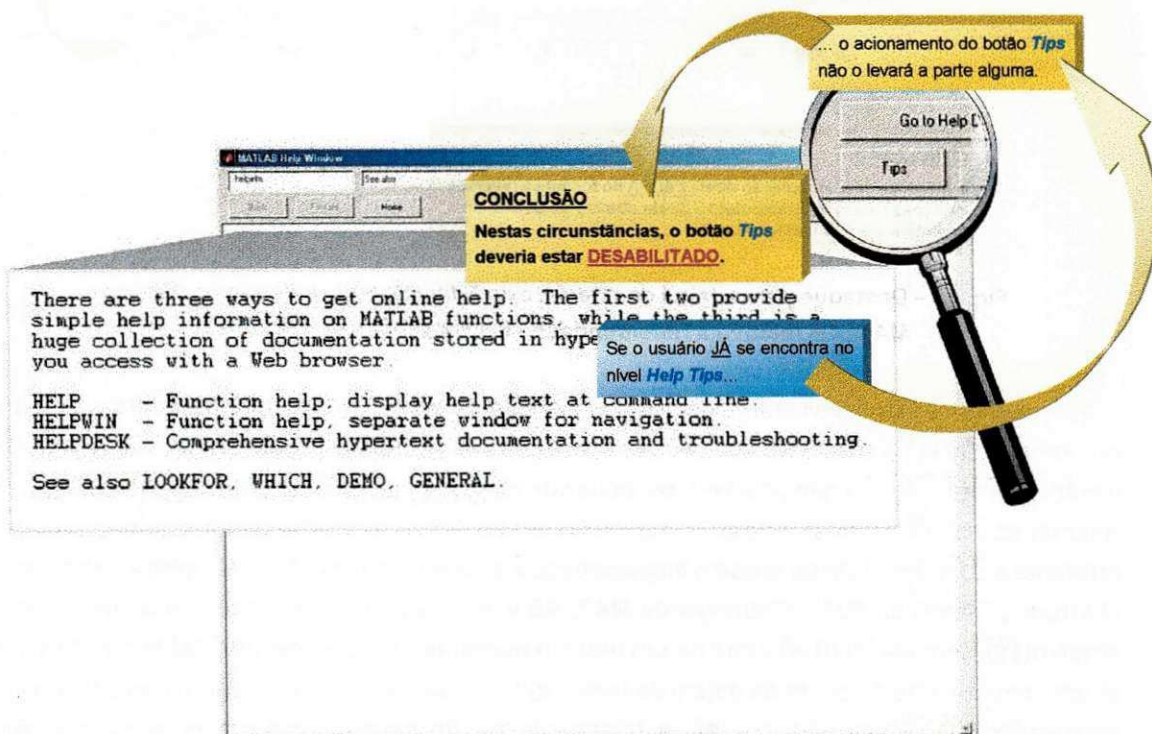


Fig. 41 – Destaque para a falha no comportamento do botão *Tips*.

Em um primeiro nível, uma falha da mesma natureza é observada quando se seleciona a sub-opção **Help Tips**. Os botões que deveriam estar habilitados na *Janela de Ajuda* quando o usuário consulta dicas (**Help Tips**) deveriam ser apenas aqueles que lhe permitissem acessar outros níveis da ajuda, não o nível no qual este se encontra. Um botão habilitado implica a possibilidade de acionamento para algum propósito e, nestas circunstâncias, o acionamento do botão **Tips** é inócuo (para não dizer confuso), i.e., não conduzindo efetivamente a nenhum resultado, visto que o usuário já se encontra no nível de informação para o qual o botão, ao ser acionado, deveria conduzi-lo, conforme destacado na Fig. 41.

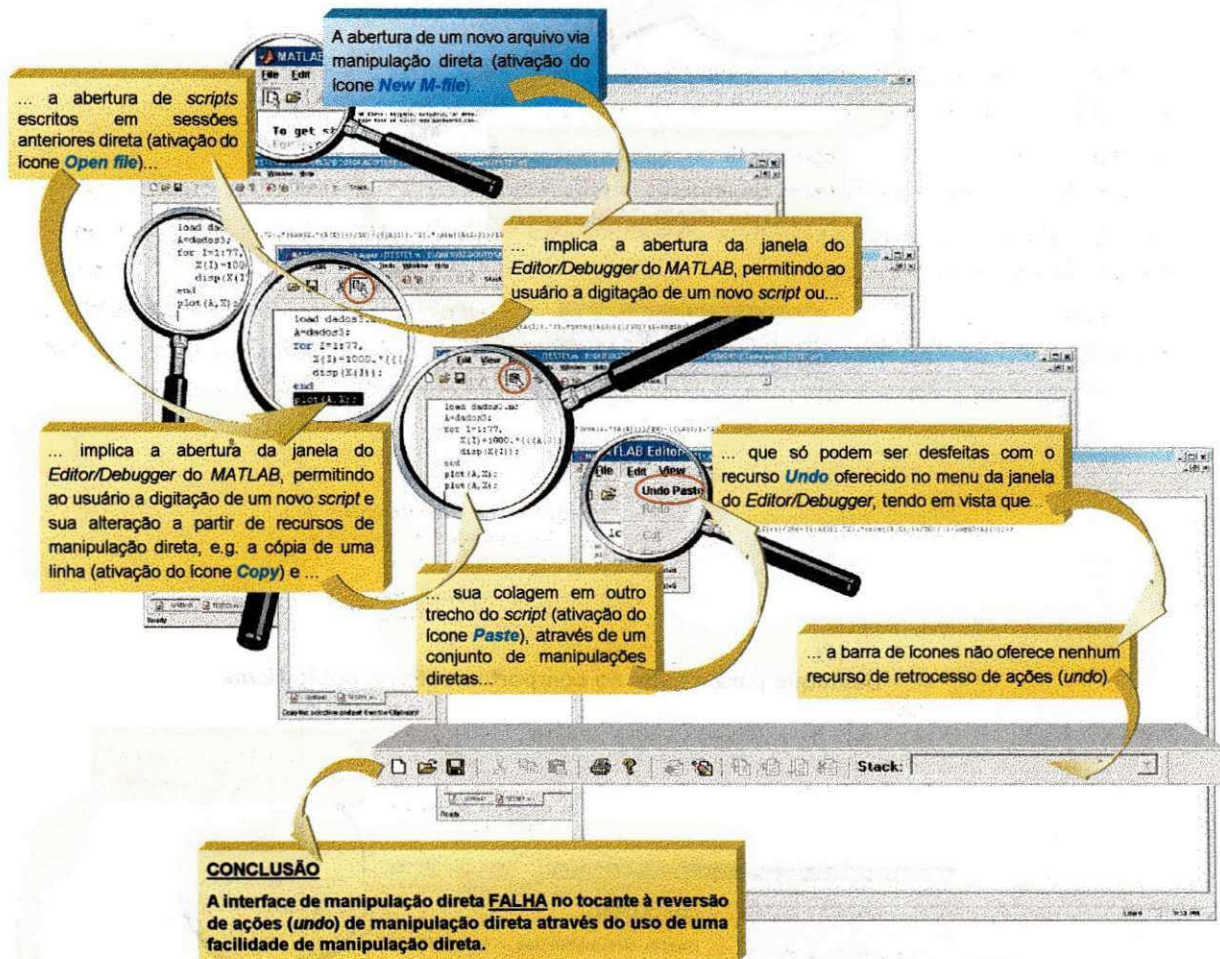







Fig. 42 – Destaque para a falha da interface do **Editor/Debugger** da versão 5.3.0 do **MATLAB** relativa à recomendação 6.1.5 da **ISO 9241-16**.

A quarta falha explicitada através da inspeção de conformidade à **ISO 9241-16** é relativa à recomendação 6.1.5 (**Retorno a estados anteriores a manipulações diretas**), conforme a qual o usuário deverá, sempre que possível, ser capaz de utilizar recursos de manipulação direta a fim de retornar ao estado anterior a uma manipulação direta. A Fig. 42 ilustra detalhadamente a falha referente a conformidade do produto inspecionado a esta recomendação, exemplificando-a com a abertura da janela do **Editor/Debugger** do **MATLAB** v. 5.3.0 a partir do ícone de criação de um novo arquivo (D), seguida da (i) abertura de um arquivo existente, (ii) cópia de uma das linhas do *script* aberto com o ícone de cópia de objeto de texto (C), (iii) da duplicação do objeto copiado a partir da ativação do ícone de colagem (P) e, finalmente, (iv) da tentativa de desfazer as manipulações diretas anteriores com outra manipulação direta. O retorno ao estado anterior só é possível a partir

da sub-opção **Undo**, acessível através da seleção da opção **Edit** do menu principal da janela do **Editor/Debugger**.

A operação de arrasto de qualquer bloco apresentado na lista de blocos da janela **Simulink Library Browser** para qualquer janela aberta a partir da ativação do ícone  (**Create a new model**) faz com que o bloco arrastado seja criado dentro janela. Esta operação é informada visualmente para o usuário através de alteração do cursor de uma seta  para uma seta apontada para um cubo .

No processo de inspeção da conformidade do produto à recomendação **6.3.3 (Diferenciação semântica no processo de arrasto)**, constatou-se, no entanto, que mesmo não havendo nenhum arquivo **.mdl**⁹² aberto, se um bloco selecionado for arrastado para qualquer área da tela o cursor  (*seleção e apontamento*) se altera para a forma  (*criação de blocos em janelas de arquivos .mdl*), embora nenhum bloco seja criado no ponto em que o botão esquerdo do *mouse* é liberado. Este comportamento pode confundir o usuário, sobretudo se for principiante, dando-lhe a falsa idéia de que a ação de arrasto para qualquer região da tela (não restrita apenas à área da janela do modelo) poderá ser executada satisfatoriamente. A Fig. 43 ilustra em detalhes a referida falha, que também está vinculada à recomendação **6.3.1 (Visualização do arrasto)**, que recomenda a indicação visual de onde é ou não possível “depositar” o objeto arrastado.

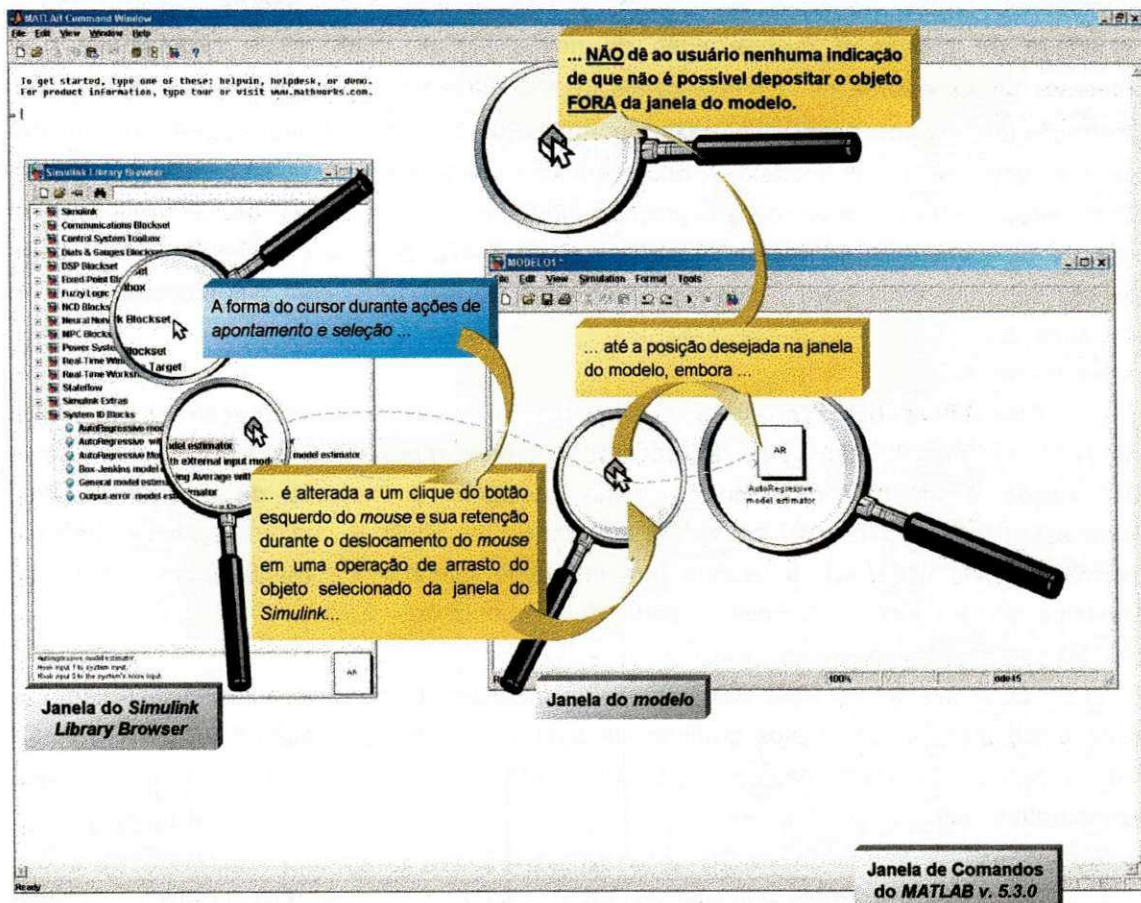


Fig. 43 – Destaque para a falha da interface do *Simulink* relativa ao arrasto de objetos.

⁹² Extensão adotada pelo *Simulink* para arquivos do tipo *modelo*.

A sexta falha detectada concerne à recomendação **6.4.2 (Mecanismos de dimensionamento)**, aplicável a contextos de trabalho que requerem o dimensionamento de objetos gráficos. Nestes casos, recomenda-se que o sistema ofereça simultaneamente mecanismos de redimensionamento unidimensional e bidimensional.

Sendo uma ferramenta empregada na modelagem, simulação e análise de sistemas dinâmicos, o *MATLAB* v. 5.3.0 interage com o *Simulink*, um módulo que dá suporte à modelagem de sistemas lineares e não lineares, à modelagem temporal contínua, amostrada ou híbrida dos dois. No *Simulink*, o usuário acessa, a partir de uma GUI e operações de seleção e arrasto, as facilidades oferecidas para a construção de modelos sob a forma de diagramas de blocos. Esta metáfora de lápis e papel, que possibilita ao usuário a realização de suas tarefas de modelagem como se o fizesse de modo convencional, serve de base para todas as operações de manipulação direta permitidas por este módulo.

De acordo com a documentação *online* (disponível em HTML) acessada a partir da sub-opção **Help Desk** da opção **Help** do *MATLAB* v. 5.3.0, tais características representam uma inovação que diferencia significativamente esta versão do *Simulink* das anteriores, que requeriam a formulação de equações diferenciais em uma linguagem ou programa. A documentação *online* também menciona (i) a existência de uma biblioteca abrangente de sumidouros, fontes, componentes lineares e não lineares e conectores, assim como a facilidade de criação e customização de novos blocos, conforme a necessidade e o interesse do usuário; e (ii) a possibilidade de modelagem por refinamento (*top-down*) e por abstração (*bottom-up*), dada a natureza hierárquica do processo de modelagem no *Simulink*. Após a definição de um modelo, a documentação *online* sugere sua simulação a partir de uma seleção de métodos de integração, diretamente dos menus do *Simulink* ou via entrada de comandos, na janela de comandos do *MATLAB*. Então, a simulação resultante poderá ser transferida para o espaço de trabalho do *MATLAB*, para pós-processamento e visualização.

Esta breve descrição teve o propósito de destacar a integração dos módulos *MATLAB* e *Simulink*, bem como de procurar esclarecer que além das manipulações diretas de criação e abertura de arquivos; cópia, colagem e exclusão de objetos gráficos; retrocesso de ações (*undo*); acesso a facilidades da ajuda e customização de gráficos, permitidas pelo *MATLAB*, o usuário também pode interagir via manipulação direta com modelos construídos no *Simulink* a partir de manipulações diretas.

Uma das ações mais elementares de manipulação direta é o dimensionamento (e redimensionamento) de objetos gráficos, da qual trata a recomendação **6.4.2** da *ISO 9241-16*. A falha identificada na versão 5.3.0 do *MATLAB* é destacada a partir do estudo comparativo ilustrado na Fig. 44.

Observando a janela principal do *MS Word2000*, verifica-se que os objetos gráficos podem ser redimensionados unidimensional (*largura* ou *altura*) e bidimensionalmente (*largura* e *altura* simultaneamente), através das alças de redimensionamento vertical ou horizontal e diagonal, respectivamente. Por outro lado, a observação da janela de comandos da versão

5.3.0 do *MATLAB* mostra que embora os objetos gráficos gerados no *Simulink* possam ser redimensionados em ambas as dimensões, tal só ocorre simultaneamente, através das alças de redimensionamento diagonal, NÃO havendo possibilidade de redimensioná-los unidimensionalmente, dada à inexistência de alças de redimensionamento vertical ou horizontal. Pode-se concluir, portanto, que o aplicativo falha no tocante à recomendação 6.4.2 (Mecanismos de dimensionamento) da *ISO 9241-16*.

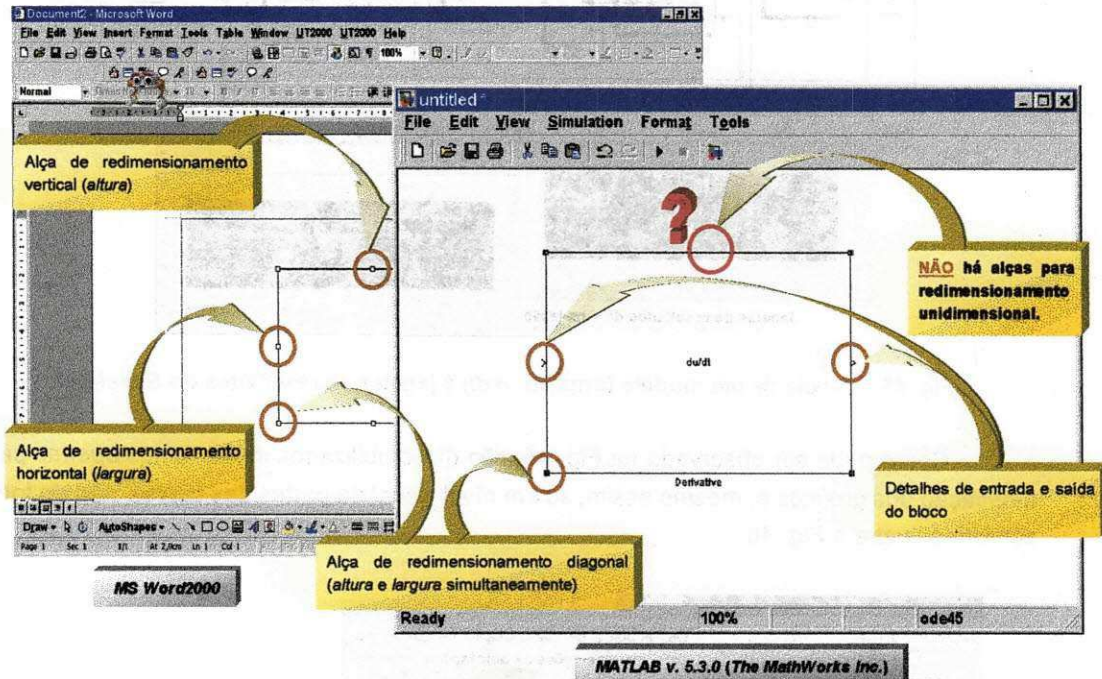


Fig. 44 – Redimensionamento de objetos gráficos no *MS Word2000* e no *MATLAB v. 5.3.0*.

A recomendação 6.4.5 (Manipulação de escala) da *ISO 9241-16* postula que, se for apropriado à tarefa, deverá ser permitida ao usuário a alteração da escala dos objetos visualizados, seja para proporcionar um grau de detalhamento mais adequado às necessidades do usuário ou para que este possa manipular a representação desses objetos mediante um único passo do processo.

Como complemento desta idéia, a declaração 6.4.6 (Manipulação direta do fator de escala) recomenda que se a tarefa requerer alterações na escala dos objetos visualizados, deverão ser oferecidos ao usuário mecanismos para a manipulação direta do fator de escala que indiquem, a partir de um movimento suave e contínuo, a alteração das dimensões do objeto, se fazendo corresponder à manipulação direta executada sobre o objeto pelo usuário.

Na versão 5.3.0 do ambiente *MATLAB*, a visualização dos resultados gráficos de uma simulação ocorre em janelas separadas do modelo considerado, como ilustra a Fig. 45.

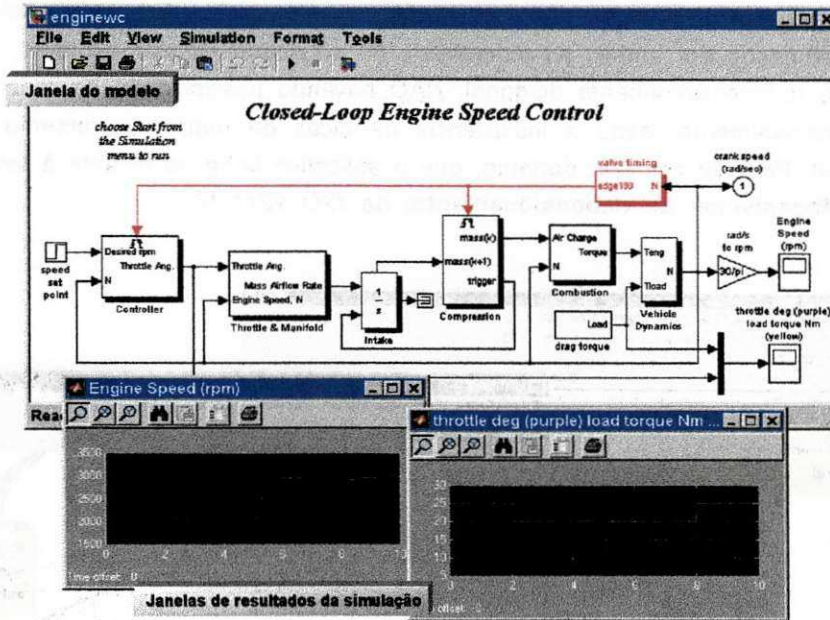


Fig. 45 – Janela de um modelo (arquivo .mdl) e janelas de resultados do Simulink.

Como pode ser observado na Fig. 45, são disponibilizados mecanismos *apenas* para ampliação dos gráficos e, mesmo assim, só em nível visual de partes das curvas, não do todo, conforme ilustra a Fig. 46.

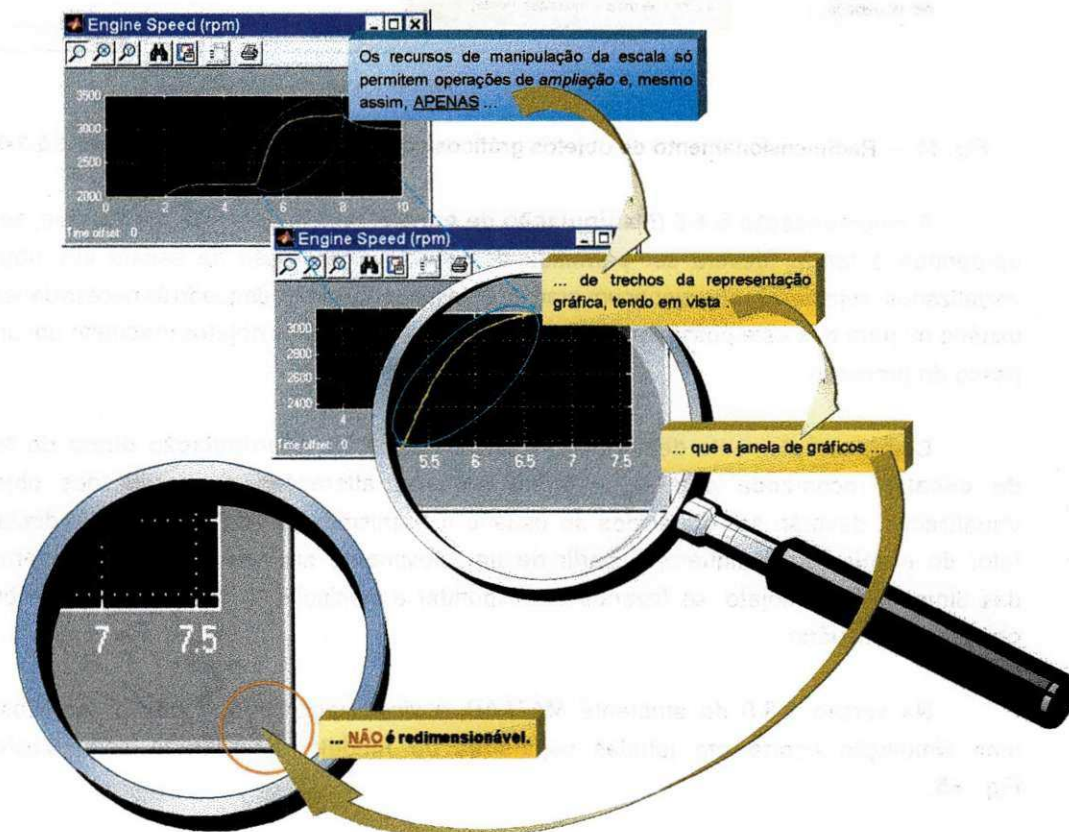


Fig. 46 – Detalhe da ampliação de um resultado gráfico em uma janela de resultados do Simulink.

NÃO são oferecidos mecanismos para ampliação ou redução para impressão do conteúdo da janela de resultados, o que pode ser verificado observando-se a ampliação da barra de ferramentas da referida janela na Fig. 47. Diante do exposto, pode-se afirmar que o aplicativo falha tanto no tocante à recomendação 6.4.5 (Manipulação de escala) quanto à recomendação 6.4.6 (Manipulação direta do fator de escala) da ISO 9241-16.

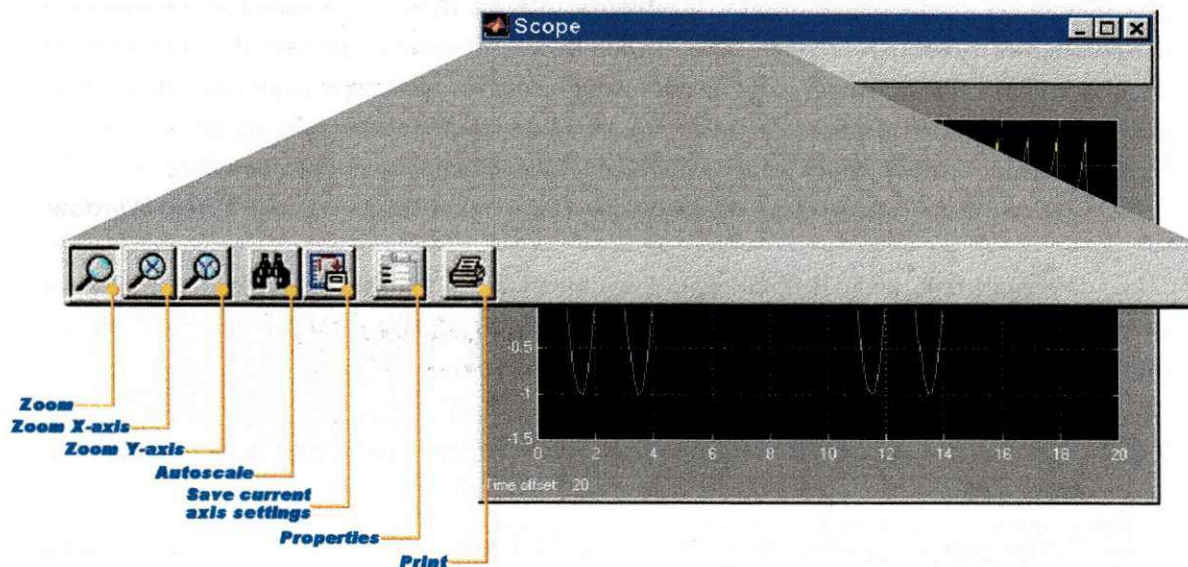


Fig. 47 – Barra de ferramentas da janela de resultados do *Simulink*.

Se a Fig. 45 for novamente observada, verificar-se-á que NÃO há nenhum mecanismo de manipulação da escala dos objetos gráficos contidos na janela do modelo, o que implica uma falha em relação à recomendação 6.4.5. Adicionalmente, também NÃO é oferecido nenhum mecanismo para a manipulação direta do fator de escala que indique suave e continuamente a alteração das dimensões de objetos, em correspondência com a manipulação direta executada sobre estes, o que constitui uma falha na conformidade do aplicativo recomendação 6.4.5.

Além destes aspectos e fugindo do escopo da ISO 9241-16, há ainda a questão relativa à ausência de mecanismos de manipulação direta da escala dos resultados gráficos de simulações, para fins de impressão, detalhada através das Figs. 46 e 47, a qual também constitui uma falha do aplicativo.

Finalmente, a última falha identificada a partir do processo de inspeção de conformidade do produto à Parte 16 do ISO 9241 relaciona-se com a declaração 8.2.1 (Rearranjo do conteúdo visualizado de uma janela de acordo com a seleção do usuário), através da qual se recomenda o movimento/rolamento automático da tela até a finalização do processo de seleção de objetos pelo usuário via manipulação direta, quando a área que os contém se estende além da borda da janela ativa.

Embora o aplicativo “passe” com relação a esta recomendação no que diz respeito ao rearranjo dos objetos de texto visualizados tanto na janela de comandos (*MATLAB Command Window*) quanto na janela do *Editor/Debugger*, de acordo com a seleção do

usuário, este falha no tocante ao movimento/rolamento automático da tela quando do rearranjo de objetos gráficos na janela de modelos do *Simulink*.

Em resumo, o processo de inspeção da conformidade do *MATLAB v.5.3.0* à *ISO 9241-16* resultou na identificação de uma dezena falhas no processo interativo, das quais se afiguram mais graves aquelas referentes às recomendações **5.4.2 (Técnicas equivalentes via teclado)**, **6.1.2 (Seqüência de manipulação direta de objetos)** e **6.1.5 (Retorno a estados anteriores a manipulações diretas)**. A primeira delas, devido à impossibilidade de acesso do usuário à barra de ferramentas caso o dispositivo de apontamento e seleção apresentar algum defeito. As duas últimas, por refletirem falhas detectadas nos demais modos de interação do produto. A não conformidade à recomendação **6.1.2** da *ISO 9241-16* se repete nos demais modos de interação, que utilizam a mesma janela para a apresentação das informações associadas à *Help Window*.

Por outro lado, a não conformidade à recomendação **6.1.5 (Retorno a estados anteriores a manipulações diretas)** da *ISO 9241-16* reflete a falha do produto no tocante à sub-cláusula **a) (Reversão)** da recomendação **6.8 (Comandos destrutivos)**.

7.2.3 Taxas de Adoção e Parecer sobre o Produto com base no Processo de Inspeção

A *ISO* recomenda, após a realização da inspeção de conformidade de um produto a um de seus padrões internacionais ou a algumas de suas partes, que os resultados da inspeção sejam sumariados a partir da computação de um indicador denominado *taxa de adoção*⁹³, definido como a razão percentual do número de recomendações julgadas satisfatoriamente adotadas pelo produto (i.e., o número de células assinaladas na coluna **P** da lista de inspeção) pelo número de recomendações julgadas aplicáveis ao contexto do projeto (i.e., o número de células assinaladas na coluna **S** da lista de inspeção).

Tabela 1 – Taxas de Adoção do *MATLAB v. 5.3.0* às Partes 14, 15 e 16 do *ISO 9241*.

PARTE DO <i>ISO 9241</i>	#P	#S	TA (%)
14	64	68	94,12
15	55	70	78,57
16	51	60	85,00

Além das taxas de adoção do *MATLAB v. 5.3.0* às Partes 14, 15 e 16 do padrão internacional *ISO 9241*, apresentadas na coluna **TA** (abreviação do termo *taxa de adoção*) da Tabela 1, são também reportados os números de células assinaladas nas colunas **P** e **S** das listas de inspeção, apresentados, respectivamente, nas colunas **#P** e **#S**, conforme

⁹³ *Adherence rating*, na versão inglesa, abreviado **AR** pela *ISO*. Em nível deste documento, interpreta-se o processo de inspeção de conformidade de um *projeto* a um padrão internacional como sendo a verificação da aplicabilidade das recomendações contidas no padrão ao contexto do projeto. Por outro lado, o processo de inspeção de conformidade de um *produto* a um padrão internacional, conforme realizado nesta pesquisa, consiste na verificação da adoção das recomendações aplicáveis ao contexto do projeto cuja implementação originou o produto inspecionado. Eis porque preferiu-se o termo *adoção* a outros termos equivalentes (e.g., *adesão*, *aderência*).

recomendado na sub-seção A.7.2 (Summary data) de cada parte considerada.

De acordo com a Tabela 1, verifica-se que o modo de interação oferecido pelo *MATLAB* v. 5.3.0 mais conforme (TA = 94,12%) às recomendações da ISO é o que envolve diálogos via *menus*, seguido do modo de interação via *manipulação direta* (TA = 85%). Em contrapartida, o modo de interação fundamentado em diálogos via *linguagem de comandos* é o que adota o menor número de recomendações da ISO aplicáveis aos contextos de uso do universo amostral considerado, apresentando, portanto, a taxa de adoção mais baixa (TA = 78,57%).

Neste ponto, vale ressaltar que a *taxa de adoção* deve ser interpretada apenas como um indicador numérico recomendado pela ISO para sumariar os resultados de um processo de inspeção de conformidade eminentemente subjetivo, que se fundamenta em julgamentos de um especialista ou de uma equipe de especialistas à luz de coletâneas de recomendações sobre um aspecto específico de *hardware* ou de *software*.

Também é bom levar em conta o fato de que essas coletâneas de recomendações em que os julgamentos do especialista têm como base princípios e regras de projeto atualmente consagradas, originadas da observação e da experimentação de pesquisadores atuantes em diferentes áreas afins da interação homem-máquina (e.g., fisiologia humana, computação gráfica, psicologia, linguística, semiótica, etnografia). Por conseguinte, não se deve deixar de considerar o cunho de subjetividade que lhes é característico. Além do mais, é importante não esquecer que a computação da *taxa de adoção* não leva em conta a ponderação das diferentes recomendações *per se*, face ao contexto de uso avaliado.

Portanto, não é recomendável encarar este indicador quantitativo como uma medida confiável do grau de conformidade às recomendações aplicáveis, sem contextualizá-lo com uma análise da importância das recomendações julgadas aplicáveis ao contexto de uso avaliado. Por outro lado, não se pode deixar de considerar que a análise da aplicabilidade de cada recomendação das *Partes 14, 15 e 16* fundamentou-se em informações de projeto repassadas ao consumidor através da documentação fornecida conjuntamente ao produto, o que implica falar de informações essencialmente necessárias à caracterização do produto adquirido. Sem dúvida, teria sido bem mais fácil de analisar a aplicabilidade dessas recomendações se se tivesse acessado as informações de concepção e desenvolvimento do produto, disponível apenas à equipe de desenvolvimento do *MATLAB*.

Eis porque nas três seções anteriores, tentou-se analisar o mais detalhadamente possível a severidade de cada falha identificada no contexto global de uso do produto avaliado, a fim de que se pudesse formular um parecer que auxiliasse a equipe de desenvolvimento do produto a otimizá-lo.

Encerrando esta seção, o Quadro 40 sumaria os problemas identificados a partir da inspeção de conformidade do *MATLAB* v. 5.3.0 às *Partes 14, 15 e 16* do padrão ISO 9241 e apresenta um parecer que reflete o ponto de vista deste enfoque avaliatório.

Quadro 40 – Sumário de Falhas e Parecer sobre o **MATLAB v. 5.3.0** com base na **Inspeção de Conformidade** com as **Partes 14, 15 e 16** do **ISO 9241**.

PARTE	# DA FALHA	COMENTÁRIO
14	01	Não há distinção entre objetos <i>não selecionáveis</i> (<i>títulos e textos explicativos</i>) e objetos <i>selecionáveis</i> (<i>itens consultáveis</i>) nas modalidades de ajuda <i>Help Windows</i> e <i>Help Tips</i> .
	02	Os designadores numéricos da lista de janelas abertas (opção <i>Window</i> do menu principal) são iniciados por 0, não por 1.
	03	O ciclo de opções do menu principal é interrompido na transição da última opção (<i>Help</i>) para a primeira (<i>File</i>) e vice-versa.
	04	Não há nenhuma tecla de atalho para mover o cursor entre grupos de opções (e.g. entre o grupo <i>Undo</i> e o grupo <i>Cut/Copy/Paste/Clear</i> da opção <i>Edit</i> do menu principal).
15	05	Não são oferecidos grupos alternativos de nomes de comandos como opções para diferentes categorias de usuários.
	06	Várias palavras de comando excedem o limite de 7 caracteres estabelecido no padrão (há palavras com até 19 caracteres).
	07	Não é permitido o uso de abreviações de nomes de comandos durante a digitação de frases de comando.
	08	Não foi adotada uma regra simples para a abreviação dos nomes de comandos.
	09	Também não foi adotado o truncamento dos nomes de comandos, mas um misto de nomes completos e truncamento (e.g., <i>cls_sys proptable, pagesetupdlg</i>).
	10	Não há nenhum mecanismo que torne óbvio para o usuário o uso de <i>teclas de funções</i> e <i>de atalhos</i> no diálogo via linha de comando.
	11	Não é permitida a redigitação ou edição <u>apenas</u> de trechos incorretos de comandos e parâmetros associados.
	12	O sistema não interpreta nem aceita a entrada de comandos contendo erros de digitação e/ou ortografia, mesmo que não haja ambigüidade quanto ao comando pretendido.
	13	Não há mecanismos para o processamento de comandos destrutivos (e.g., <i>delete, clear</i>).
	14	Não há mecanismos para a reversão de ações do usuário (<i>undo</i>).
	15	O sistema executa <u>qualquer</u> sentença de comando correta <u>sem</u> solicitar antes nenhuma confirmação do usuário.
16	O sistema não oferece nenhuma realimentação sobre o tempo restante para a conclusão de processos em andamento.	
17	O sistema não evidencia os erros cometidos em linhas de comando (as mensagens de erro são lacônicas e pouco esclarecedoras).	
16	18	Não há mecanismo de acesso à barra de ferramentas de manipulação direta via teclado (dispositivo de entrada alternativo do <i>mouse</i>)
	19	Os botões <i>Home</i> e <i>Tips</i> nas janelas de ajuda <i>Help Windows</i> e <i>Help Tips</i> contrariam a seqüência de manipulação direta de objetos, estando habilitados em circunstâncias em que não o deveriam.
	20	A barra de ferramentas de manipulações diretas do <i>Editor/Debugger</i> não possui ícone de retrocesso de ações (<i>undo</i>), o que impede o retorno do usuário a estados anteriores.
	21	Não há indicação visual de onde um objeto gráfico do <i>Simulink</i> pode ser "depositado" pelo usuário no processo de criação de um modelo (arquivo <i>.mdl</i>).
	22	Não há diferenciação semântica no cursor durante o processo de arrasto de um objeto gráfico da janela do <i>Simulink</i> para a janela do modelo, o que pode confundir o usuário.
	23	O sistema não permite o dimensionamento horizontal ou vertical de objetos gráficos, <u>apenas</u> o redimensionamento horizontal e vertical simultâneo de objetos "depositados" na janela do modelo.
	24	O sistema não oferece mecanismos para a ampliação ou redução do conteúdo das janelas de resultados para fins de impressão.
	25	O sistema não oferece mecanismos para a manipulação direta do fator de escala que, suave e continuamente, indiquem a alteração das dimensões de objetos gráficos.
	26	O sistema não oferece mecanismos de movimento/rolamento automático da tela quando do rearranjo de objetos gráficos na janela de modelos do <i>Simulink</i> .
PARECER:		Com base nas Partes 14, 15 e 16 do padrão ISO 9241 , embora o produto ofereça ao usuário recursos poderosos para a solução de diversos contextos de problemas, há diferentes categorias de falhas em sua interface que necessitam ser corrigidas, a fim de que o processo de interação usuário-produto seja otimizado, conforme atesta a lista acima apresentada.

7.3 Resultados da Mensuração do Desempenho

Esta seção apresenta, em duas sub-seções, os resultados obtidos a partir do enfoque avaliatório *mensuração do desempenho* do usuário durante o uso do aplicativo avaliado, o *MATLAB v. 5.3.0*.

Na primeira sub-seção, serão apresentados os indicadores quantitativos coletados a partir do ensaio de usabilidade conduzido no *Laboratório de Interfaces Homem-Máquina (LIHM)* do *DEE/CCT/UFPB* para a avaliação do *MATLAB v. 5.3.0*, fundamentada na *observação direta com registro em vídeo* do uso do produto por **40** usuários de teste, assim como na *verbalização informal de ações (informal think aloud)* pelos usuários de teste. Nesta seção também será feita uma análise informal dos resultados, contextualizada pelos eventos observados durante as sessões de teste, assim como por análises retrospectivas dos registros em vídeo dessas sessões.

A segunda sub-seção apresentará os problemas identificados a partir do processo de observação direta respaldado pelas análises retrospectivas dos registros em vídeo das sessões de teste.

7.3.1 Análise Preliminar dos Indicadores Quantitativos

A Tabela 2 apresenta os indicadores quantitativos coletados através de um processo de *observação direta* que contou com a participação de um grupo de professores e estudantes de pós-graduação (doutorado e mestrado) e graduação dos cursos de *Engenharia Elétrica*, *Engenharia Civil*, *Engenharia Mecânica* e *Bacharelado em Matemática* do *Campus II* do *Centro de Ciências e Tecnologia* da *Universidade Federal da Paraíba*.






Conforme discutido no capítulo anterior, após a etapa de recrutamento dos usuários de teste e a devolução dos questionários *pré-teste*, o universo amostral voluntário totalizava **58** participantes, tendo sido subdividido em três sub-grupos – *princiantes*, *intermediários* e *experientes*, conforme a experiência com o produto e a frequência de uso. No entanto, apenas **74%** do contingente considerado, ou seja **43** indivíduos, compareceram às sessões de testes destinadas à avaliação do desempenho e da satisfação dos usuários com o produto-alvo.

Adicionalmente, os resultados obtidos nos testes de aproximadamente **5%** dos participantes se mostraram inaceitáveis para o propósito da pesquisa, por questões essencialmente relativas à experiência, habilidades e predisposição ao uso do produto pelos participantes, motivo pelo qual foram excluídos do experimento.

Ao final deste processo de triagem, o universo amostral efetivo deste experimento passou a totalizar **40** indivíduos, de ambos os sexos, com predominância do sexo masculino, distribuídos em três categorias, à proporção de **16:16:8**, como se pode verificar observando a Tabela 2.






Tabela 2 – Indicadores Quantitativos Obtidos a partir do Teste de Usabilidade (Observação Direta com Registro em Vídeo)

SÍNTESE DOS INDICADORES QUANTITATIVOS REGISTRADOS

LEGENDA																																					
		 Tempo de Execução da Tarefa	 Número de Ações Incorretas	 Número de Opções Incorretas	 Número de Erros Repetidos	 Número de Consultas à Ajuda																															
Usuário	Tarefa	U0	U1	U2	U3	U4	U5	U6	U7	U8	U9	U10	U11	U12	U13	U14	U15	U16	U17	U18	U19	U20	U21	U22	U23	U24	U25	U26	U27	U28	U29	U30	U31	U32	U33	U34	U35
T0	153	202	275	279	238	276	254	NCC	220	298	195	276	242	297	281	291	268	204	293	212	232	168	195	262	259	211	NCC	247	280	119	296	216	150	142	238	143	
	2	0	2	2	0	1	2	18	1	4	1	1	0	7	2	2	2	0	0	1	12	2	1	4	1	1	5	0	3	0	2	1	0	0	2	0	
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
	0	0	0	1	0	0	0	10	0	0	0	0	0	4	0	0	1	0	0	0	10	1	0	3	0	0	2	0	1	0	0	0	0	0	0	0	
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T1	303	312	598	598	596	NCC	432	325	NCC	303	428	451	475	359	306	471	NCC	306	318	305	416	351	353	402	359	475	579	320	413	447	NCC	393	344	293	278	312	
	4	0	4	6	4	8	5	10	8	1	6	11	1	12	9	2	12	0	2	1	6	2	1	0	2	3	2	0	2	0	3	1	1	2	1	0	
	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	0	0	0	6	0	0	4	0	0	0	0	8	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
	0	0	0	0	2	0	0	0	2	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
T2	421	NCC	NCC	NCC	NCC	NCC	NCC	NCC	NCC	598	562	NCC	NCC	NCC	NCC	NCC	NCC	NCC	597	594	597	563	NCC	NCC	534	NCC	NCC	NCC	NCC	593	596	593	NCC	469	282	598	
	14	13	6	6	16	0	8	12	11	6	10	5	4	7	4	5	6	4	4	9	3	4	9	11	4	11	8	8	6	2	10	12	0	1	0	0	
	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	5	2	2	3	14	2	4	8	7	4	5	1	2	5	0	6	4	1	1	4	0	2	4	5	2	4	4	3	4	0	4	3	0	0	1	0	
	2	0	3	0	0	0	2	0	0	0	0	0	0	1	0	0	0	2	0	0	0	0	2	0	1	0	0	0	0	1	0	0	0	0	0	0	0
T3	265	249	NCC	539	258	597	NCC	494	490	496	124	196	598	NCC	600	566	269	359	225	248	174	159	236	381	276	345	572	593	321	471	NCC	342	384	260	597	324	
	4	1	10	10	1	6	12	11	3	5	2	1	1	5	6	2	3	0	2	0	1	1	3	2	8	4	0	0	0	1	10	6	11	2	0	2	
	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0	0	4	4	0	3	6	4	2	4	0	0	0	3	3	1	1	0	0	0	0	0	0	0	0	5	0	0	0	0	6	3	6	0	0	1	
	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0
T4	537	597	596	NCC	NCC	NCC	NCC	NCC	NCC	NCC	599	NCC	NCC	NCC	NCC	NCC	NCC	NCC	NCC	NCC	NCC	NCC	517	506	NCC	NCC	NCC	NCC	NCC	NCC	495	NCC	543	429	264	NCC	396
	9	11	3	12	3	1	9	12	0	10	12	15	3	6	12	5	6	2	9	0	2	2	4	9	4	0	5	4	5	3	5	4	1	0	3	2	
	0	3	12	3	2	0	8	9	8	4	3	3	0	5	3	7	2	1	0	4	3	2	0	5	4	5	3	3	7	2	0	1	0	0	0	0	
	8	10	9	1	1	1	6	14	4	9	9	1	2	2	5	4	1	1	5	1	0	1	1	4	5	2	4	1	4	3	5	3	0	0	0	1	
	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0	1	0	2	0	0	2	0	0	0	0	0	2	0

A segunda linha da Tabela 2 contém uma legenda com os símbolos atribuídos aos cinco indicadores objetivos considerados, que aparecem na última coluna relacionados aos dados coletados em cada uma das tarefas do teste de usabilidade conduzido no *LIHM/DEE/CCT/UFPB*. Em virtude do volume de dados e, por conseguinte, das dimensões da Tabela 2, o Quadro 41 apresenta os referidos símbolos em dimensões que possibilitam uma definição melhor dos detalhes, assim como as abreviações empregadas em algumas das tabelas deste capítulo.

Quadro 41 – Indicadores Quantitativos Considerados e Símbolos Correlatos

INDICADOR	SÍMBOLO ADOTADO
Tempo de Execução da Tarefa (T_e)	
Número de Ações Incorretas (N_{ai})	
Número de Opções Incorretas (N_{oi})	
Número de Erros Repetidos (N_{er})	
Número de Consultas à Ajuda (N_{ca})	






Na segunda coluna da Tabela 2 (**U0**), destacada das demais pela cor da letra e da moldura (vermelho escuro), assim como pela cor de preenchimento das células (ciano), se encontram os resultados obtidos a partir da realização do teste-piloto. Embora não inclusos no processamento estatístico, estes resultados foram inseridos na Tabela 2 para fins de comparação com aqueles referentes às diferentes categorias (*principiante*, *intermediário* e *experiente*) em que foi particionado o universo amostral de usuários de teste, destacadas entre si pelo uso de diferentes tons de amarelo no preenchimento das células correspondentes.

Outra convenção adotada na Tabela 2 diz respeito à desistência ou ao não encerramento de uma tarefa em tempo hábil pelo usuário de teste. Observa-se, pois, que algumas células das linhas de registro do tempo de execução das tarefas foram preenchidas com a representação **NCC**, utilizada para os casos em o usuário de teste **Não Conseguiu Concluir** a tarefa, seja por desistência ou pelo não encerramento em tempo hábil⁹⁴.

A fim de facilitar a interpretação dos dados coletados, construiu-se uma versão sumariada da Tabela 2, a Tabela 3, que contém apenas os valores mínimos e máximos dos indicadores quantitativos considerados, para cada categoria pré-definida de usuários de teste, onde o *tempo de execução* é apresentado em segundos (s).

⁹⁴ Conforme pode ser observado nos Anexos I1 a I5, os tempos de execução das tarefas de teste foram de 5 minutos (300 s), para a Tarefa 0 (Pré-Tarefa) e 10 minutos (600 s) para as Tarefas 1 a 4.

Tabela 3 – Número de indivíduos associados aos mínimos nulos registrados em função das categorias pré-definidas de usuários de teste.

INDICADORES		TAREFAS											
		T1			T2			T3			T4		
		Valor	P	I	E	P	I	E	P	I	E	P	I
	mínimo	303	305	278	552	534	282	124	159	198	596	429	264
	máximo	598	579	418	598	597	597	600	593	597	599	543	498
	mínimo	0	0	0	0	0	0	1	0	0	0	0	0
	máximo	12	6	2	16	12	4	12	11	2	15	9	5
	mínimo	0	0	0	0	0	0	0	0	0	0	0	0
	máximo	5	0	0	1	0	0	0	0	0	12	7	4
	mínimo	0	0	0	0	0	0	0	0	0	0	0	0
	máximo	8	2	0	14	5	2	6	6	1	14	5	4
	mínimo	0	0	0	0	0	0	0	0	0	0	0	0
	máximo	2	0	0	3	2	0	1	3	0	1	4	2

NOTA: O indicador *tempo de execução* está representado em *segundos* (s).

A **Tarefa 0** (*Pré-Tarefa*) foi suprimida da Tabela 3, por objetivar apenas a verificação da familiaridade do usuário de teste com sistemas computacionais e, mais especificamente, com o ambiente *MS Windows* (inicialização do sistema, criação de pasta de trabalho, transferência de arquivos pertinentes ao contexto das tarefas para a pasta de trabalho e abertura do aplicativo sob condições de teste). As colunas **P** (*principiantes*), **I** (*intermediários*) e **E** (*experientes*) mantêm a consistência cromática da Tabela 2.

É necessário acrescentar, entretanto, que os dados referentes à **Tarefa 0** foram incluídos no processamento estatístico, cujos resultados serão apresentados e discutidos na próxima seção. O propósito da inclusão foi verificar se havia diferenças estatisticamente identificáveis nas ações das três categorias de usuários de teste consideradas, no que diz respeito à inicialização do sistema, entrada com senha no ambiente *Windows*, criação de uma pasta de trabalho, transferência de um grupo de arquivos e ativação do *MATLAB v. 5.3.0*.

Grosso modo, uma análise descontextualizada da Tabela 3, em especial das linhas relativas aos *tempos de execução* mínimos e máximos, poderia conduzir à conclusão de que os resultados obtidos para as três categorias de usuários de teste consideradas não apresentam diferenças notáveis. Entretanto, tal conclusão seria precipitada, além de falsa.

Em primeiro lugar, por não se tratar de um universo amostral equidistribuído, i.e., todas as classes com igual número de indivíduos. É conveniente recordar que neste experimento as categorias de usuários de teste foram distribuídas à proporção de 16:16:8, ou seja, 2:2:1. Em segundo lugar, porque a Tabela 3 apresenta *apenas* os valores mínimos (e máximos) dos indicadores quantitativos considerados, não fornecendo nenhuma informação sobre o número de usuários de teste de cada categoria associados aos mínimos verificados. Finalmente, porque a análise de um conjunto de valores é sempre incompleta quando o contexto segundo o qual os dados foram obtidos não é convenientemente considerado.

Considere-se inicialmente o indicador *tempo de execução* ao longo das diferentes tarefas de teste. Observa-se que não há diferenças significativas nos mínimos e máximos obtidos para as diferentes categorias pré-definidas de usuários de teste. No entanto, ao integrar estes resultados aos máximos referentes aos indicadores *número de ações incorretas*, *número de opções incorretas* e *número de erros repetidos*, as diferenças entre categorias começam a se tornar evidentes. Se tal integração contar com a informação sobre o número de usuários de teste de cada categoria associados aos mínimos nulos registrados, então as diferenças serão ainda mais evidentes.

A Tabela 4 complementa as informações apresentadas na Tabela 3, possibilitando a verificação, por categoria de usuários de teste e por tarefa, do número de indivíduos para os quais foram registrados os mínimos nulos dos indicadores *número de ações incorretas*, *número de opções incorretas*, *número de erros repetidos* e *número de consultas à ajuda*. Tendo em vista tratar-se de uma amostra de pequenas dimensões, recomenda-se estatisticamente a apresentação de valores percentuais em segundo plano, razão pela qual aparecem entre parênteses, porém destacados dos demais pela cor da fonte (vermelho escuro), a fim de facilitar a visualização.

Tabela 4 – Número de indivíduos associados aos mínimos nulos registrados em função das categorias pré-definidas de usuários de teste.

		NÚMERO DE USUÁRIOS ASSOCIADOS AOS MÍNIMOS NULOS											
Categoria		T1			T2			T3			T4		
Indicador		P	I	E	P	I	E	P	I	E	P	I	E
		1 (6,25)	1 (6,25)	5 (62,50)	1 (6,25)	1 (6,25)	3 (37,50)	0 (0,00)	5 (31,25)	5 (62,50)	1 (6,25)	2 (12,50)	1 (12,50)
		15 (93,75)	16 (100,00)	8 (100,00)	15 (93,75)	16 (100,00)	8 (100,00)	16 (100,00)	16 (100,00)	8 (100,00)	2 (12,50)	4 (25,00)	4 (50,00)
		12 (75,00)	15 (93,75)	8 (100,00)	1 (6,25)	3 (18,75)	6 (75,00)	5 (31,25)	12 (75,00)	7 (87,50)	0 (0,00)	2 (12,50)	5 (62,50)
		12 (75,00)	16 (100,00)	8 (100,00)	13 (81,25)	12 (75,00)	8 (100,00)	14 (87,50)	15 (93,75)	8 (100,00)	15 (93,75)	12 (75,00)	5 (62,50)

O processo de observação direta, assim como a investigação retrospectiva do registro em vídeo das sessões de teste, permitiu uma análise contextualizada do indicador *tempo de execução* que explica o porquê das diferenças pouco significativas nos valores registrados. Na verdade, os mínimos e máximos registrados para cada categoria pré-definida são próximos entre si por diferentes razões. Inicialmente, deve-se considerar que, ao início de cada sessão, o usuário de teste foi instruído a executar a tarefa como se o fizesse em seu ambiente de trabalho, sem uma preocupação excessiva com o tempo de execução. O propósito desta instrução foi tentar atenuar a pressão imposta ao usuário pela delimitação do tempo, assim como o caráter intrusivo da observação direta e, em especial, do registro em vídeo.

Além disso, é necessário lembrar que se buscou elevar progressivamente o grau de dificuldade das ações do usuário, de tarefa a tarefa, quer em termos de complexidade (e.g., a transição da simples digitação de um conjunto de dados para a identificação de erros de lógica em um *script*), quer a partir da solicitação de resultados que exigiam o uso de recursos novos ou pouco conhecidos do produto avaliado (e.g., a configuração customizada das propriedades das linhas, dos textos e da perspectiva de apresentação de uma representação gráfica).

Evidentemente, tanto as reações individuais ao ambiente controlado e todas as suas implicações (e.g., delimitação do tempo de execução, ações condicionadas a roteiros de teste) quanto as posturas dos participantes face ao grau de experiência com sistemas computacionais e, sobretudo, com o produto avaliado, implicaram diferenças observáveis no uso do tempo delimitado para a execução de cada tarefa, em função de seu grau de dificuldade. Por conseguinte, a observação direta e a análise retrospectiva dos registros em vídeo possibilitaram o delineamento de diferentes padrões de ações realizadas ao longo dos intervalos de tempo estipulados que, grosso modo, caracterizaram as diferentes categorias de usuários de teste consideradas.

De um modo geral, as sessões com os usuários *principiantes* foram caracterizadas por um dispêndio de grande parte dos tempos pré-definidos para a execução das tarefas de teste com a:

- (i) identificação dos recursos adequados dentre aqueles disponibilizados pelo produto avaliado;
- (ii) interpretação (nem sempre correta) das solicitações dos roteiros de teste;
- (iii) análise (nem sempre satisfatória) das respostas do produto;
- (iv) tentativa (muitas vezes infrutífera) de reversão de ações ou opções incorretas.

No uso dos recursos de áudio para a comunicação entre o usuário de teste e os avaliadores, predominaram, do lado dos usuários, os questionamentos sobre:

- (i) recursos mais adequados a situações específicas (e.g., o uso do acessório *WordPad* do *MS Windows* ou do *Editor/Debugger* do *MATLAB v. 5.3.0* para a criação dos arquivos de dados e edição dos dados);
- (ii) procedimentos necessários para o acesso ou a configuração de determinados recursos (e.g., o comando para o carregamento de um arquivo de dados na área de trabalho, a opção para a alteração da cor de uma linha em um gráfico);
- (iii) estratégias apropriadas à reversão de situações de erro;
- (iv) formas mais eficientes de obtenção de determinados resultados (e.g., a configuração das propriedades das curvas de uma representação gráfica).

Nessas sessões, os avaliadores representaram não apenas os papéis de observadores, instrutores e mediadores, mas também de tutores, dadas às constantes solicitações de ajuda dos usuários de teste. Foi praticamente inexistente a *verbalização espontânea de ações (think aloud)*, solicitada informalmente no início de cada sessão de teste, o que exigiu uma atenção redobrada dos avaliadores às ações do usuário e ao grau de interferência no processo, face às solicitações de ajuda, assim como um maior número de análises retrospectivas dos registros em vídeo, para melhor entendimento e verificação mais acurada dos problemas identificados ao longo das sessões de teste.

Por outro lado, o dispêndio de grande parte dos tempos pré-definidos para a execução das tarefas de teste ao longo das sessões com os usuários *experientes* ocorreu devido à:

- (i) explanação detalhada e sequencial das ações associadas a cada tarefa de teste;
- (ii) análise silenciosa (via de regra correta) das solicitações dos roteiros de teste;
- (iii) ponderação e discussão de diferentes estratégias possíveis para a execução das tarefas;
- (iv) análise explicativa das respostas do produto.

No uso dos recursos de áudio para a comunicação entre o usuário de teste e os avaliadores, predominou, do lado dos usuários, a verbalização (espontânea ou induzida) de ações no tocante a:

- (i) vantagens e desvantagens inerentes ao uso de diferentes modos de interação (e.g., o

número de etapas necessárias para a criação dos arquivos de dados e edição dos dados a partir do uso do *WordPad* do *MS Windows* ou do *Editor/Debugger* do *MATLAB v. 5.3.0*);

(ii) questionamentos sobre detalhes das solicitações dos roteiros de teste (e.g., a necessidade da digitação de um determinado volume de dados em um arquivo);

(iii) aspectos referentes ao acesso e configuração de determinadas facilidades (e.g., a configuração das propriedades das curvas de uma representação gráfica via programação ou a partir de manipulação direta);

(iv) formas mais eficientes de obtenção de determinados resultados (e.g., otimização de um *loop* em um *script* de teste).

Nas sessões com os usuários *experientes*, os avaliadores atuaram essencialmente como observadores e, esporadicamente, como instrutores e mediadores, dada à familiaridade dos usuários de teste com os recursos oferecidos pelo produto avaliado, bem como com suas limitações. A *verbalização de ações (think aloud)*, *espontânea* ou *induzida pelo avaliador*, marcou a maior parte das sessões de teste, permitindo aos avaliadores a formulação de questionamentos adicionais das ações menos compreensíveis dos usuários de teste.

Um fato comum às sessões de teste da maioria dos usuários *principiantes* e de alguns usuários *experientes* consistiu na necessidade de um maior número de análises retrospectivas dos registros em vídeo para melhor entendimento de ações associadas à execução de determinadas tarefas, embora por razões completamente distintas. O maior número de análises retrospectivas dos registros das sessões com os usuários *principiantes* se deu, na maior parte dos casos, para a verificação de seqüências confusas de ações ou para o entendimento de grupos de erros várias vezes repetidos, enquanto no caso de alguns dos usuários *experientes* as análises retrospectivas foram várias vezes repetidas visando o acompanhamento de determinadas seqüências de ações, dada a rapidez com que foram executadas.

No que diz respeito aos usuários *intermediários*, as sessões foram caracterizadas por aspectos comuns a ambos os extremos, i.e., *principiantes* e *experientes*. O perfil do usuário *intermediário* deste experimento retrata, em essência, um indivíduo cuja familiaridade com os recursos oferecidos pelo produto avaliado é um pouco superior àquela exibida por um usuário *principiante* característico, embora não se equipare (raras as exceções) à de um usuário *experiente*.

As *verbalizações de ações (think aloud)* durante as sessões ocorreram (exceto para quatro indivíduos) muito mais *induzidas pelo avaliador* do que *espontaneamente* e com teores de informação mais restritos quando comparadas àquelas feitas pelos usuários *experientes*. Via de regra, o dispêndio de grande parte dos tempos pré-definidos para a execução das tarefas de teste ao longo das sessões com os usuários *intermediários* deveu-se à reversão de situações de erro, na maior parte dos casos em virtude da falta de atenção à leitura dos roteiros de teste, assim como da ponderação insuficiente das alternativas e da precipitação na execução das ações, em geral visando a abreviação da conclusão das tarefas.

Retornando à Tabela 3 e aos riscos de se chegar a conclusões enganosas a partir de análises superficiais e descontextualizadas dos dados, tome-se como exemplo os tempos de

execução mínimos relativos à **Tarefa 3** (*Alteração de dados numéricos*), na qual foi solicitada ao usuário a correção de uma descontinuidade no gráfico resultante da execução do *script* de teste. Observando-se as células correspondentes aos tempos de execução mínimos das três categorias de usuários de teste, verifica-se que o menor dos mínimos (**124 s**) é relativo à coluna dos usuários *principiantes*, enquanto o maior destes (**198 s**) se encontra registrado na coluna dos usuários *experientes*, o que conduz à conclusão incorreta de que em termos do encerramento da tarefa no menor intervalo de tempo os usuários *principiantes* se mostraram mais ágeis do que os *experientes*.

No entanto, o processo de observação direta, aliado às análises retrospectivas dos registros em vídeo das sessões de teste, demonstra que esta conclusão é falsa, à medida que esclarece tal inversão: os usuários *principiantes* (menor tempo mínimo) simplesmente substituíram ou excluíram da série de dados o valor da abscissa que provocava a descontinuidade da função, enquanto alguns usuários *intermediários* e a maioria dos *experientes* dispenderam mais tempo corrigindo a descontinuidade a partir da interpolação de seus vizinhos imediatos e, por conseguinte, produzindo resultados matematicamente mais aceitáveis.

Diante do exposto, verifica-se que a inclusão das informações apresentadas na Tabela 4 e dos aspectos que caracterizaram as sessões de teste com as diferentes categorias de usuários enriquece o contexto da análise dos valores mínimos e máximos dos indicadores quantitativos considerados, fornecendo elementos adicionais à interpretação superficial do início desta discussão, que fundamentam melhor a explicação da aparente similaridade dos dados apresentados na Tabela 3, em especial no que se refere ao *tempo de execução*.

No tocante ao *número de ações incorretas* e ao *número de erros repetidos*, observa-se na Tabela 4, principalmente para os limites superiores, que à medida que aumenta o grau de dificuldade das tarefas aumenta a diferença entre os resultados associados a cada categoria de usuários de teste, sendo confirmada as expectativas de melhores desempenhos dos usuários *experientes*, seguidos dos *intermediários* que, por sua vez, apresentam desempenho superiores àqueles exibidos pelos *principiantes*.

O indicador *número de opções incorretas* refere-se mais especificamente às opções oferecidas pelo modo de interação via menus. Vale a pena salientar que, no tocante ao modo de interação do usuário com o aplicativo avaliado, facultou-se aos usuários a escolha daquele mais adequado ao seu grau de familiaridade com o aplicativo, suas habilidades com os diferentes dispositivos típicos de entrada de dados (teclado e *mouse*) e, evidentemente, suas preferências individuais.

Os processos de observação direta e análise retrospectiva dos registros em vídeo mostraram que os usuários *experientes* utilizaram predominantemente o modo de interação via linguagem de comandos para a realização das tarefas de teste envolvendo a entrada e manipulação de dados (**Tarefas 1 a 3**), enquanto que as demais categorias alternaram o uso dos diálogos via linguagem de comandos com diálogos via menus, tendo sido observada e registrada a preferência pelo uso de menus.

Quanto à **Tarefa 4**, dada à natureza das ações, observou-se a predominância do uso de diálogos via menus e manipulação direta por todas as categorias de usuários de teste. A resistência ao uso destes modos de interação por alguns usuários propiciou o cometimento de um número maior de ações incorretas, em virtude do grau de complexidade do uso de recursos de linguagem de comandos para a customização de representações gráficas ser superior à complexidade de uso dos recursos equivalentes nos demais modos de interação, no *MATLAB v. 5.3.0*.

De um modo geral, as categorias de usuários *intermediários* e *principiantes* demonstraram maior familiaridade no uso de recursos disponibilizados via menus e manipulação direta do que nos mecanismos de interação via linguagem de comandos. Devido provavelmente à familiaridade com aplicativos de uso geral (editores de texto, planilhas eletrônicas, editores gráficos) dotados de interfaces gráficas, tal fato se reflete na Tabela 3, nas linhas referentes ao *número de ações incorretas* e *número de opções incorretas*. Uma análise comparativa dos limites superiores dos indicadores acima citados mostra que os valores associados ao *número de ações incorretas* são superiores àqueles relativos ao *número de opções incorretas*, ao longo de todas as tarefas de teste, para as categorias de usuários *intermediários* e *principiantes*.

Embora tal análise comparativa não esclareça adequadamente este fato, a observação direta e as análises retrospectivas mostraram que foram cometidas mais ações incorretas do que feitas opções incorretas, uma vez que os usuários de teste apresentaram melhor desempenho no uso dos recursos oferecidos pelo diálogo via menus do que na utilização dos mecanismos de interação via linguagem de comandos e manipulação direta, não porque o uso dos diálogos via linguagem de comandos e manipulação direta foi mais efetivo do que o uso de diálogos via menus.

O indicador *número de consultas à ajuda*, conforme apresentado nas Tabelas 2 e 3, só possibilita a conclusão de que os usuários de teste recorreram *muito pouco* aos mecanismos de ajuda *offline* e *online* desenvolvidos pela *The MathWorks Inc.* para a versão 5.3.0 do *MATLAB*, i.e., o guia do usuário (*User's Guide* [Hans97]), a janela de ajuda (*Help Window*), as dicas (*Help Tips*), os tutoriais (*Help Desk*) e os exemplos e demonstrações (*Examples and Demos*), todos disponíveis aos usuários de teste durante a realização das sessões de avaliação.

Entretanto, o processo de observação e, principalmente, as análises retrospectivas dos registros em vídeo permitiram constatar que das 35 consultas realizadas ao longo de todas as sessões de teste⁹⁵, incluindo o teste-piloto, apenas 8 (22,85%) consultas foram plenamente satisfatórias, i.e., os usuários conseguiram contornar completamente, a partir das consultas à ajuda, as dificuldades enfrentadas na execução das tarefas propostas.

Adicionalmente, segundo as verbalizações de ações (espontâneas e/ou induzidas pelos avaliadores) registradas durante as sessões e incluídas informalmente neste experimento, 25% do universo amostral - 7 (17,5%) usuários *principiantes* e 3 (7,5%) usuários *intermediários* - declararam ter dificuldade na leitura da língua inglesa, enquanto 17 (40%) dos participantes consultaram os avaliadores sobre o significado de pelo menos uma palavra inglesa, quer durante às consultas a uma das formas de ajuda acima mencionadas, quer durante as tentativas de interpretação das mensagens de erro do aplicativo avaliado.

⁹⁵ O que corresponde à taxa de menos de 1 consulta por usuário.

7.3.2 Análise Estatística dos Indicadores Quantitativos

Diversos experimentos científicos possibilitam a aquisição de dados de natureza subjetiva e objetiva, similarmemente ao processo de observação com registro em vídeo conduzido em nível do ensaio de usabilidade realizado no âmbito desta pesquisa. Em geral, dados de natureza subjetiva se mostram valiosos para a contextualização de dados numéricos, como se pôde verificar ao longo da análise preliminar feita na seção anterior.

As anotações referentes aos indicadores objetivos considerados⁹⁶, aliadas à verbalização de ações pelos usuários de teste e ao registro em vídeo das sessões, possibilitaram não só a contextualização dos dados quantitativos coletados, como também a caracterização das sessões de teste relativas a cada sub-grupo de usuários que compôs o universo amostral do experimento. Todavia, apesar de elucidar diversos questionamentos, a referida análise preliminar *per si* não dá margem a que se faça alguma espécie de generalização, uma vez que envolve aspectos subjetivos que dizem respeito apenas ao universo amostral estudado.

Antes de tudo, é necessário salientar que a análise da eficácia da observação direta, com registro em vídeo, na identificação de falhas na interface do *MATLAB v. 5.3.0* foi um dos propósitos da condução do ensaio de usabilidade. Visou-se também o planejamento de um experimento laboratorial destinado à medição de indicadores objetivos de usabilidade em uma amostra representativa de usuários do produto⁹⁷, cujos resultados, após uma análise estatística adequada, permitissem inferências confiáveis para o todo.

Conforme a revisão apresentada no Anexo M, não é possível fazer estatísticas sem amostras aleatórias das populações de interesse. No entanto, a amostragem aleatória das populações de interesse tem se mostrado, em diversas situações do cotidiano, impraticável ou provavelmente impossível, segundo Edgington [Edgi95], que acrescentou que praticamente todos os experimentos não permitem inferências estatísticas válidas sobre as populações de interesse. Na impossibilidade da amostragem aleatória, as inferências estatísticas sobre os efeitos de tratamento deveriam ser restritas aos indivíduos (ou outras unidades experimentais) envolvidos no experimento, enquanto as inferências sobre os efeitos de tratamento para outros indivíduos (ou outras unidades experimentais) deveriam ser denominadas *inferências não estatísticas*, i.e. inferências sem base em probabilidades, uma vez que se generaliza as unidades experimentais a outras que são bastante semelhantes no tocante às características consideradas pertinentes [Edgi95, Sieg97].

Não há neste comentário nenhum intuito de questionar ou criticar os procedimentos adotados no âmbito da Estatística Clássica à luz de diversos autores (e.g. Edgington [Edgi95], Siegel [Sieg97]) que se preocupam com os problemas de validade das estatísticas padrões em experimentos do mundo real.

No entanto, não se pode deixar de considerar questões relativas à validade do processamento estatístico dos dados coletados a partir do ensaio de usabilidade conduzido no

⁹⁶ Vide Anexo L (modelos das fichas de registro de eventos utilizadas nas sessões de teste).

⁹⁷ É conveniente não esquecer que, embora a avaliação tenha sido conduzida em nível da interface do produto e não do produto como um todo, a interface É o produto, para a maioria dos usuários de produtos de *software*.

âmbito desta pesquisa, face às dificuldades de cunho prático enfrentadas no processo de amostragem, principalmente quanto ao número disponível de usuários do produto, aos custos adicionais necessários à ampliação da abrangência geográfica e à predisposição dos usuários disponíveis à participação não remunerada.

Tais restrições implicaram a necessidade inevitável do uso de experiência e julgamento não estatísticos desde o processo de recrutamento dos participantes. Esta situação se enquadra perfeitamente no comentário de Edginton [Edgi95] sobre o principal "fardo" do pesquisador no processo de generalização a partir de experimentos: as inferências têm sido e continuam a ser feitas a partir de lógica não estatística, ao invés de lógica estatística.

A aplicação de análise estatística de dados experimentais obtidos de amostras, como é o caso do ensaio de usabilidade conduzido no âmbito deste trabalho de pesquisa, é bastante relevante, pois possibilita a generalização, com determinado nível de confiança, dos resultados obtidos a partir da amostra considerada.

No tocante à seleção do universo amostral do experimento ora discutido, foi feito um levantamento dos usuários do *MATLAB* no *Campus II* da *UFPB*, através do qual identificou-se 286 usuários. Devido a grande parte da população ser composta por estudantes de cursos de graduação (*Engenharia Elétrica, Engenharia Civil, Engenharia Mecânica e Bacharelado em Matemática*), o recrutamento em sala de aula ou laboratório foi bastante difícil. Além do mais, constatou-se entre os indivíduos recrutados um baixo índice de predisposição à participação no experimento. Apesar disso, ainda foi possível distribuir 148 questionários *pré-teste*, dos quais foi possível recolher 58 exemplares. Deste total de respondentes, apenas 43 compareceram às sessões de teste, dos quais 3 foram excluídos, por razões já apresentadas.

Conquanto a amostra não tenha podido ser selecionada de forma aleatória, dadas as circunstâncias acima expostas (o que evitaria a tendência dos resultados), seus indivíduos (adultos de ambos os sexos, professores e estudantes universitários da *UFPB*) apresentaram características bastantes distintas, como será discutido na próxima seção, o que possibilitou a composição de um universo amostral representativo da população considerada de usuários do *MATLAB*. O tamanho da amostra selecionada excedeu 30 unidades amostrais, de modo que se pode categorizá-la como *grande amostra* ($n_A > 30$), segundo a literatura [Burt96, Sabo99, Fech00].

Os valores obtidos para os indicadores quantitativos pré-definidos se encontram organizados na Tabela 2, apresentada e analisada na sub-seção anterior. O propósito principal da análise estatística foi o de encontrar evidências suficientes para concluir que havia um efeito de tratamento significativo nos fatores de interesse (*indicadores objetivos medidos*) ao longo dos grupos considerados (*usuários principiantes, intermediários e experientes*), tendo em vista tratar-se de um experimento ao longo do qual três categorias de usuários de teste executaram cinco (05) tarefas sequenciais e com graus de dificuldade crescentes, em cada uma das quais foram registrados os valores relativos aos cinco (05) indicadores objetivos.

Conforme a revisão de estatística apresentada no Anexo M e, mais especificamente de acordo com a sub-seção M3 do referido anexo, o procedimento estatístico mais adequado ao propósito almejado foi o teste *F* ANOVA fator único. Uma vez que este teste evidencia apenas diferenças entre médias, não possibilitando comparações entre pares de grupos, utilizou-se adicionalmente para este propósito o procedimento *Tukey-Kramer*.

No início das sessões de teste, o participantes foram instruídos a encerrarem cada tarefa no instante indicado pelo avaliador, a fim de respeitar os *tempos de execução* pré-definidos nos roteiros das tarefas (300 s para a **Tarefa 0** e 600 s para as demais). No entanto, previu-se um "intervalo de prorrogação" equivalente a 50% dos tempos pré-definidos para a execução das tarefas (i.e. 150 s para a **Tarefa 0** e 300 s para as demais), que deveria ser concedido pelo avaliador ao participante, caso os eventos ocorridos ao longo dos tempos pré-definidos indicassem a possibilidade de identificação de falhas.

Evidentemente, tal estratégia foi adotada apenas para efeito de identificação de falhas, não para a medição do indicador *tempo de execução*. Nos casos em que o "intervalo de prorrogação" foi utilizado, o avaliador utilizou a abreviação **NCC** (**N**ão **C**onseguiu **C**oncluir) no registro do tempo de execução, embora tenha continuado a observar as ações do usuário de teste até o final do "intervalo de prorrogação", independentemente do usuário de teste haver ou não concluído a tarefa durante o referido intervalo. Assim, **NCC** deve ser interpretado como **Não Conseguiu Concluir em tempo hábil** ou, simplesmente, **Não Conseguiu Concluir**, o que para efeito de medição do *tempo de execução* tem o mesmo significado.

O processamento estatístico dos dados foi realizado com o auxílio do *MS Excel2000*, que consistiu das etapas descritas a seguir. Inicialmente, gerou-se um relatório de estatísticas univariadas para os dados coletados, cujo sumário⁹⁸ é apresentado na Tabela 5. Em seguida, construiu-se as matrizes de correlação dos indicadores objetivos pré-definidos, apresentada na Tabela 6, com o propósito de analisar a força das relações entre as variáveis consideradas.

Como se pode observar na Tabela 6, omitiu-se alguns dos elementos das matrizes de correlação, em virtude de indeterminações nos cálculos dos coeficientes de correlação associados a linhas da Tabela 2 que contêm séries de dados nulos, e.g. linhas 6 (*número de opções incorretas*) e 8 (*número de consultas à ajuda*).

Por outro lado, destacou-se com cor de fundo azul os coeficientes de correlação relativos a fortes associações ($0,75 < \rho \leq 1$ para correlações *positivas* ou $-0,75 > \rho \geq -1$ para correlações *negativas*) entre os indicadores objetivos medidos durante as sessões de teste do ensaio de usabilidade.

⁹⁸ O sumário a que se refere o texto consiste de um sub-conjunto do relatório de estatísticas univariadas do *MS Excel2000*, que inclui, além daquelas apresentadas na Tabela 5, outras estatísticas irrelevantes a esta análise (*moda, erro padrão, curtose, assimetria, intervalo, máximo, mínimo, soma, contagem, maior, menor e nível de confiança*).

Tabela 5 – Sumário da análise descritiva dos dados coletados.

SÍNTESE DE ESTATÍSTICAS UNIVARIADAS																
Tarefa	Estatística	PRINCIPIANTES					INTERMEDIÁRIOS					EXPERIENTES				
		T _e (s)	N _{ai}	N _{oi}	N _{er}	N _{ca}	T _e (s)	N _{ai}	N _{oi}	N _{er}	N _{ca}	T _e (s)	N _{ai}	N _{oi}	N _{er}	N _{ca}
T0	Média	259,47	2,81	0	1,00	0	222,93	2,06	0	1,06	0	204,88	0,88	0	0,13	0
	Desvio Padrão	33,17	4,40	0	2,61	0	51,81	3,04	0	2,54	0	59,40	1,46	0	0,35	0
	Variância	1100,41	19,36	0	6,80	0	2684,35	9,26	0	6,46	0	3527,84	2,13	0	0,13	0
T1	Média	434,92	6,19	0,31	1,38	0,38	385,40	1,88	0	0,13	0	337,29	0,63	0	0	0
	Desvio Padrão	111,56	4,00	1,25	2,60	0,72	74,43	2,03	0	0,50	0	56,47	0,92	0	0	0
	Variância	12444,91	16,03	1,56	6,78	0,52	5540,54	4,12	0	0,25	0	3188,57	0,84	0	0	0
T2	Média	575,00	7,44	0,06	4,44	0,38	583,38	6,56	0	2,56	0,38	480,86	2,00	0	0,50	0
	Desvio Padrão	32,53	4,03	0,25	3,33	0,89	22,93	3,65	0	1,71	0,72	144,31	1,93	0	0,76	0
	Variância	1058,00	16,26	0,06	11,06	0,78	525,98	13,33	0	2,93	0,52	20824,48	3,71	0	0,57	0
T3	Média	421,23	4,94	0	2,19	0,13	339,07	3,06	0	1,25	0,19	342,75	0,75	0	0,13	0
	Desvio Padrão	173,95	3,89	0	1,94	0,34	129,91	3,71	0	2,32	0,75	124,01	1,04	0	0,35	0
	Variância	30258,36	15,13	0	3,76	0,12	16876,21	13,80	0	5,40	0,56	15377,93	1,07	0	0,13	0
T4	Média	597,33	7,50	4,50	4,94	0,06	498,00	3,81	2,50	2,50	0,56	388,25	2,50	1,38	1	0,75
	Desvio Padrão	1,53	4,68	3,39	4,12	0,25	42,49	2,61	2,13	1,83	1,15	95,90	1,77	1,69	1,60	1,04
	Variância	2,33	21,87	11,47	17,00	0,06	1805,00	6,83	4,53	3,33	1,33	9196,25	3,14	2,84	2,57	1,07

Tabela 6 – Matrizes de correlação.

ANÁLISE DA CORRELAÇÃO ENTRE OS INDICADORES QUANTITATIVOS																
Tarefa	Indicador	PRINCIPIANTES					INTERMEDIÁRIOS					EXPERIENTES				
		Te (s)	Nai	Noi	Ner	Nca	Te (s)	Nai	Noi	Ner	Nca	Te (s)	Nai	Noi	Ner	Nca
T0	Te (s)	1,000000					1,000000					1,000000				
	Nai	0,638896	1,000000				0,204192	1,000000				0,707618	1,000000			
	Noi			1,000000					1,000000					1,000000		
	Ner	0,363500	0,970264		1,000000		0,110560	0,964539		1,000000		0,626715	0,866199		1,000000	
	Nca					1,000000					1,000000					1,000000
T1	Te (s)	1,000000					1,000000					1,000000				
	Nai	-0,147908	1,000000				0,304032	1,000000				-0,554665	1,000000			
	Noi		0,120723	1,000000					1,000000					1,000000		
	Ner	-0,204488	0,561820	0,473542	1,000000			0,237508		1,000000					1,000000	
	Nca	0,393025	0,112933	-0,139122	0,133541	1,000000					1,000000					1,000000
T2	Te (s)	1,000000					1,000000					1,000000				
	Nai	-1,000000	1,000000				0,288004	1,000000				0,651301	1,000000			
	Noi		0,566208	1,000000					1,000000					1,000000		
	Ner	-1,000000	0,789978	0,766679	1,000000		-0,015538	0,874214		1,000000		0,154969	0,392232		1,000000	
	Nca		-0,067710	-0,112987	-0,172683	1,000000	-0,534881	-0,187354		-0,182896	1,000000					1,000000
T3	Te (s)	1,000000					1,000000					1,000000				
	Nai	0,496508	1,000000				-0,104828	1,000000				-0,183078	1,000000			
	Noi			1,000000					1,000000					1,000000		
	Ner	0,634630	0,929452		1,000000		-0,010070	0,948111		1,000000		-0,061094	0,487950		1,000000	
	Nca	-0,101369	0,206991		0,062889	1,000000	0,540754	-0,219872		-0,143444	1,000000					1,000000
T4	Te (s)	1,000000					1,000000					1,000000				
	Nai	0,928571	1,000000				0,776257	1,000000				0,938785	1,000000			
	Noi	-0,755929	-0,063152	1,000000			0,453098	-0,065899	1,000000			0,762971	0,836899	1,000000		
	Ner	-0,906867	0,344088	0,546795	1,000000		0,671053	0,677684	0,154349	1,000000		0,053877	0,653275	0,687311	1,000000	
	Nca		-0,427699	0,275625	-0,060641	1,000000	-0,039474	0,369255	-0,095056	0,364231	1,000000		0,700649	0,470960	0,688530	1,000000

O Quadro 42 sumaria os pares de indicadores objetivos medidos que apresentaram maior tendência à covariação, traduzida pela forte correlação positiva ($\rho > 0,75$) ou negativa ($\rho < -0,75$) entre eles.

Quadro 42 – Correlações mais fortes identificadas entre pares de indicadores objetivos

PAR DE INDICADORES		SINAL DE ρ	TAREFA	SUB-GRUPO
<i>tempo de execução</i>	<i>número de ações incorretas</i>	+	0	<i>experientes</i>
			4	<i>principiantes, intermediários e experientes</i>
<i>tempo de execução</i>	<i>número de opções incorretas</i>	-	4	<i>principiantes e experientes</i>
<i>tempo de execução</i>	<i>número de erros repetidos</i>	-	4	<i>principiantes</i>
<i>número de ações incorretas</i>	<i>número de opções incorretas</i>	+	4	<i>experientes</i>
<i>número de ações incorretas</i>	<i>número de erros repetidos</i>	+	0	<i>principiantes, intermediários e experientes</i>
			2	<i>principiantes e intermediários</i>
			3	
<i>número de ações incorretas</i>	<i>número de consultas à ajuda</i>	+	4	<i>experientes</i>
<i>número de opções incorretas</i>	<i>número de erros repetidos</i>	+	2	<i>principiantes</i>

Embora destacadas com cor de fundo azul, porém contornadas em vermelho, as correlações negativas máximas determinadas entre os pares *tempo de execução* x *número de ações* e *tempo de execução* x *número de erros repetidos* para o sub-grupo *principiantes* (Tarefa 2) não foram incluídas no Quadro 42, uma vez que não refletem o desempenho da maioria dos indivíduos deste sub-grupo: apenas dois (02) usuários de teste *principiantes* conseguiram concluir a tarefa em tempo hábil (julgamento não estatístico). Por outro lado, os coeficientes de correlação que implicam uma tendência um pouco menor ($0,65 < \rho \leq 0,75$) à covariação de pares de indicadores da Tabela 6 foram destacados por um contorno azul.

Além de evidenciar tendências à covariação de pares de indicadores objetivos pré-definidos, a construção das matrizes de correlação também teve o propósito de selecionar as variáveis passíveis de serem submetidas ao teste *F ANOVA fator único* e, posteriormente, ao procedimento *Tukey-Kramer*, conforme destaca o Quadro 43.

Quadro 43 – Variáveis selecionadas para o teste *F ANOVA fator único*.

TAREFA	INDICADORES OBJETIVOS	JUSTIFICATIVA
0	<i>tempo de execução</i>	Conforme se observa nas Tabelas 2 e 5, não é possível conduzir o teste ANOVA para as variáveis N_{oi} e N_{ca} , tendo em vista pelo menos uma das séries de dados relativos às categorias <i>principiantes, intermediários e experientes</i> ser nula.
	<i>número de ações incorretas</i>	
	<i>número de erros repetidos</i>	
1	<i>tempo de execução</i>	Conforme se observa nas Tabelas 2 e 5, não é possível conduzir o teste ANOVA para as variáveis N_{oi} , N_{er} e N_{ca} , uma vez que pelo menos uma das séries de dados relativos às categorias <i>principiantes, intermediários e experientes</i> é nula.
	<i>número de ações incorretas</i>	
2	<i>tempo de execução</i>	Semelhante ao que ocorre com os dados referentes à Tarefa 0, não é possível conduzir o teste ANOVA para as variáveis N_{oi} e N_{ca} , tendo em vista pelo menos uma das séries de dados relativos às categorias <i>principiantes, intermediários e experientes</i> ser nula.
3	<i>número de ações incorretas</i>	
3	<i>tempo de execução</i>	
4	<i>tempo de execução</i>	Única tarefa cujas séries de dados permitem a condução do teste ANOVA para todas as variáveis consideradas no experimento, tendo em vista que não há nenhuma série de dados totalmente nula.
	<i>número de ações incorretas</i>	
	<i>número de opções incorretas</i>	
	<i>número de erros repetidos</i>	
	<i>número de consultas à ajuda</i>	

A análise preliminar das séries de dados da Tabela 2 visou evidenciar diferenças entre os dados coletados para as diferentes categorias de usuários de teste consideradas no experimento. Por outro lado, o propósito da análise desses dados, à luz do teste *F ANOVA fator único*, é verificar se as diferenças comentadas na seção anterior são estatisticamente aceitáveis, de modo que se possa incrementar as inferências não estatísticas com conclusões estatísticas sobre a população a partir da amostra selecionada. Assim, a *hipótese nula* de nenhuma diferença nas médias aritméticas

$$H_0: \mu_P = \mu_I = \mu_E$$

será testada em relação à *alternativa* de haver diferenças entre as referidas médias

$$H_A: \text{nem todas as médias são iguais}$$

onde μ_P , μ_I e μ_E representam as médias aritméticas relativas às séries de dados associadas, respectivamente às categorias *princiantes*, *intermediários* e *experientes*. É necessário observar que a *hipótese alternativa* não é exatamente a negação da *hipótese nula*, visto que neste experimento foram consideradas três séries de dados, não duas. Assim, o teste *F ANOVA fator único* permitiu verificar se a *hipótese nula* deveria ou não ser aceita, i.e., a interpretação a ser dada aos casos em que a *hipótese nula* foi rejeitada deve ser limitada à constatação de diferenças entre as médias. Não se pode afirmar que $\mu_P \neq \mu_I \neq \mu_E$. Por este motivo, também foi empregado nesta análise o *procedimento Tukey-Kramer*, que possibilita a determinação de médias significativamente diferentes dentre grupos de médias comparados.

Em essência, o *procedimento Tukey-Kramer* permite o exame comparativo simultâneo de todos os pares de grupos de um experimento passíveis de serem submetidos ao teste *F ANOVA fator único*. A partir do cálculo das diferenças entre todos os $c(c-1)/2$ pares de médias, seguido da obtenção de um *intervalo crítico*, determinado pela equação:

$$\text{Intervalo crítico} = Q_{S(c, n-c)} \sqrt{\frac{MQD}{2} \left(\frac{1}{n_j} + \frac{1}{n_{j'}} \right)}, \quad (8)$$

onde $Q_{S(c, n-c)}$ - valor crítico da cauda superior do intervalo de Student com c e $(n-c)$ graus de liberdade

MQD - Média dos Quadrados Dentro dos grupos

n_j - número de elementos do grupo j

$n_{j'}$ - número de elementos do grupo j' ,

compara-se cada diferença absoluta entre pares de médias ($|\bar{X}_j - \bar{X}_{j'}|$) com o valor crítico obtido. As diferenças absolutas que excederem o valor do intervalo crítico deverão ser interpretadas como resultantes de médias significativamente *diferentes*.

A Tabela 7 sumaria os resultados do teste *F ANOVA fator único* para as variáveis apresentadas no Quadro 43. Observe-se que para os testes em que a *hipótese nula* foi rejeitada os pares F e F_{crit} aparecem destacados sobre fundo amarelo claro ($F > F_{crit}$), assim como os pares α e p aparecem destacados sobre fundo alaranjado ($p < \alpha$). Por sua vez, a Tabela 8 e o Quadro 44 contêm os resultados do *procedimento Tukey-Kramer* para comparações múltiplas de médias.

Tabela 7 - Resultados da aplicação do teste F ANOVA fator único para as variáveis do Quadro 43.

		ANOVA (Single Factor) $\alpha = 0,05$																												
		Te(s)						NeI						NeII						Nca										
		Source of Variation	SS	df	MS	F	P-value	F crit	SS	df	MS	F	P-value	F crit	SS	df	MS	F	P-value	F crit	SS	df	MS	F	P-value	F crit				
T0	Between Groups	18343,8531	2	9171,9265	4,132480147	0,024479859	3,26741656	20,1250	2	10,0625	0,838069781	0,440589557	3,251926728	5,2875	2	2,6438	0,489552706	0,616814217	4,082380656	199,8125	37	5,4003								
	Within Groups	77681,5417	35	2219,4726				444,2500	37	12,0068																				
	Total	96025,3947	37					464,3750	39						205,1000	39														
T1	Between Groups	45395,0198	2	22697,5099	2,952066182	0,066590871	3,29453067	235,4375	2	117,7188	15,44194549	1,3308E-05	3,251926728	232,9375	37	6,2804														
	Within Groups	246037,9516	32	7688,6860				282,0625	37	7,6233				517,5000	39															
	Total	291432,9714	34					517,5000	39																					
T2	Between Groups	41985,7384	2	20992,8692	2,266231589	0,140399069	3,738890086	166,1250	2	83,0625	6,540702314	0,003695888	3,251926728	83,6250	2	41,8125	6,657611619	0,003380728	3,251926728	232,3750	37	6,2804								
	Within Groups	129686,7321	14	9263,3380				469,8750	37	12,6993																				
	Total	171672,4706	16					636,0000	39					316,0000	39															
T3	Between Groups	54406,14786	2	27203,0739	1,269710413	0,294279019	3,284924333	95,7250	2	47,8625	0,012262815	0,026480289	3,251926728	23,2875	2	11,6438	3,114821509	0,056204538	3,251926728	138,3125	37	3,7382								
	Within Groups	707012,741	33	21424,6285				441,3750	37	11,8291				138,3125	37	3,7382														
	Total	761418,8889	35					537,1000	39					161,8000	39															
T4	Between Groups	76124,8333	2	38062,4167	9,839934795	0,005432428	4,256492048	172,5375	2	86,2688	7,054993784	0,002537318	3,251926728	60,9000	2	30,4500	4,335353535	0,02034388	3,251926728	94,8375	2	47,4188	5,432920457	0,008536043	3,251926728					
	Within Groups	34813,4167	9	3869,1574				452,4375	37	12,2280				259,8750	37	7,0236				322,9375	37	8,7280			3,2250	2	1,6125	2,102643172	0,136490546	3,251926728
	Total	110938,2500	11					624,9750	39					320,7750	39					417,7750	39				31,6000	39				
T0	Between Groups	18343,8531	2	9171,9265	4,132480147	0,024479859	4,106482265	20,1250	2	10,0625	0,838069781	0,440589557	4,082380656	5,2875	2	2,6438	0,489552706	0,616814217	4,082380656	199,8125	37	5,4003								
	Within Groups	77681,5417	35	2219,4726				444,2500	37	12,0068																				
	Total	96025,3947	37					464,3750	39						205,1000	39														
T1	Between Groups	45395,0198	2	22697,5099	2,952066182	0,066590871	4,148773789	235,4375	2	117,7188	15,44194549	1,3308E-05	4,082380656	232,9375	37	6,2804														
	Within Groups	246037,9516	32	7688,6860				282,0625	37	7,6233				517,5000	39															
	Total	291432,9714	34					517,5000	39																					
T2	Between Groups	41985,7384	2	20992,8692	2,266231589	0,140399069	4,856710170	166,1250	2	83,0625	6,540702314	0,003695888	4,082380656	83,6250	2	41,8125	6,657611619	0,003380728	4,082380656	232,3750	37	6,2804								
	Within Groups	129686,7321	14	9263,3380				469,8750	37	12,6993																				
	Total	171672,4706	16					636,0000	39					316,0000	39															
T3	Between Groups	54406,14786	2	27203,0739	1,269710413	0,294279019	4,133767106	95,7250	2	47,8625	0,012262815	0,026480289	4,082380656	23,2875	2	11,6438	3,114821509	0,056204538	4,082380656	138,3125	37	3,7382								
	Within Groups	707012,741	33	21424,6285				441,3750	37	11,8291				138,3125	37	3,7382														
	Total	761418,8889	35					537,1000	39					161,8000	39															
T4	Between Groups	76124,8333	2	38062,4167	9,839934795	0,005432428	5,714698272	172,5375	2	86,2688	7,054993784	0,002537318	4,082380656	60,9000	2	30,4500	4,335353535	0,02034388	4,082380656	94,8375	2	47,4188	5,432920457	0,008536043	4,082380656					
	Within Groups	34813,4167	9	3869,1574				452,4375	37	12,2280				259,8750	37	7,0236				322,9375	37	8,7280			3,2250	2	1,6125	2,102643172	0,136490546	4,082380656
	Total	110938,2500	11					624,9750	39					320,7750	39					417,7750	39				31,6000	39				

Tabela 8 – Resultados da aplicação do procedimento Tukey-Kramer aos dados da Tabela 2.

COMPARAÇÕES MÚLTIPLAS DE TUKEY-KRAMER							
Tarefa		Indicadores Objetivos					
		T _e (s)	N _{ai}	N _{oi}	N _{er}	N _{ca}	
T0	Média do Grupo 1	259,4666667	2,8125			1	
	n do Grupo 1	15	16			16	
	Média do Grupo 2	222,9333333	2,0625		1,0625		
	n do Grupo 2	15	16			16	
	Média do Grupo 3	204,875	0,875		0,125		
	n do Grupo 3	8	8			8	
	MQD	2219,4726	12,0068		5,4003		
	Estatística Q	3,44	3,44		3,44		
	Comparação entre PRINCIPANTES e INTERMEDIÁRIOS						
	Diferença Absoluta	36,53333333	0,75		0,0625		
	Erro Padrão da Diferença	12,16407996	0,866269183		0,580965674		
	Intervalo Crítico	41,84443506	2,979965989		1,99852192		
	Médias dos Grupos 1 e 2 são	Não Diferentes	Não Diferentes		Não Diferentes		
	Comparação entre PRINCIPANTES e EXPERIENTES						
	Diferença Absoluta	54,59166667	1,9375		0,875		
	Erro Padrão da Diferença	14,58421953	1,060958739		0,71153473		
	Intervalo Crítico	50,16971519	3,649698062		2,447679472		
	Médias dos Grupos 1 e 3 são	Diferentes	Não Diferentes		Não Diferentes		
	Comparação entre INTERMEDIÁRIOS e EXPERIENTES						
	Diferença Absoluta	18,05833333	1,1875		0,9375		
	Erro Padrão da Diferença	14,58421953	1,060958739		0,71153473		
	Intervalo Crítico	50,16971519	3,649698062		2,447679472		
	Médias dos Grupos 2 e 3 são	Não Diferentes	Não Diferentes		Não Diferentes		
	Média do Grupo 1	434,9230769	6,1875				
	n do Grupo 1	13	16				
	Média do Grupo 2	408,1875	1,625				
	n do Grupo 2	16	16				
	Média do Grupo 3	337,2857143	0,625				
	n do Grupo 3	7	8				
	MQD	11232,20573	7,623310811				
Estatística Q	3,44	3,44					
Comparação entre PRINCIPANTES e INTERMEDIÁRIOS							
Diferença Absoluta	26,73557692	4,5625					
Erro Padrão da Diferença	27,98239343	0,690258593					
Intervalo Crítico	96,25943339	2,374489561					
Médias dos Grupos 1 e 2 são	Não Diferentes	Diferentes					
Comparação entre PRINCIPANTES e EXPERIENTES							
Diferença Absoluta	97,63736264	5,5625					
Erro Padrão da Diferença	35,13272438	0,845390672					
Intervalo Crítico	120,8565719	2,908143912					
Médias dos Grupos 1 e 3 são	Não Diferentes	Diferentes					
Comparação entre INTERMEDIÁRIOS e EXPERIENTES							
Diferença Absoluta	70,90178571	1					
Erro Padrão da Diferença	33,96037159	0,845390672					
Intervalo Crítico	116,8236783	2,908143912					
Médias dos Grupos 2 e 3 são	Não Diferentes	Não Diferentes					
Média do Grupo 1	421,2307692	4,9375		2,1875			
n do Grupo 1	13	16		16			
Média do Grupo 2	339,0666667	3,0625		1,25			
n do Grupo 2	15	16		16			
Média do Grupo 3	342,75	0,75		0,125			
n do Grupo 3	8	8		8			
MQD	21424,62852	11,92905405		3,738175676			
Estatística Q	3,44	3,44		3,7			
Comparação entre PRINCIPANTES e INTERMEDIÁRIOS							
Diferença Absoluta	82,16410256	1,875		0,9375			
Erro Padrão da Diferença	39,21961827	0,863461567		0,483359059			
Intervalo Crítico	134,9154869	2,970307792		1,788428518			
Médias dos Grupos 1 e 2 são	Não Diferentes	Não Diferentes		Não Diferentes			
Comparação entre PRINCIPANTES e EXPERIENTES							
Diferença Absoluta	78,48076923	4,1875		2,0625			
Erro Padrão da Diferença	46,50874602	1,057520126		0,591991528			
Intervalo Crítico	159,9900863	3,637869235		2,190368655			
Médias dos Grupos 1 e 3 são	Não Diferentes	Diferentes		Não Diferentes			
Comparação entre INTERMEDIÁRIOS e EXPERIENTES							
Diferença Absoluta	3,683333333	2,3125		1,125			
Erro Padrão da Diferença	45,312179	1,057520126		0,591991528			
Intervalo Crítico	155,8738958	3,637869235		2,190368655			
Médias dos Grupos 2 e 3 são	Não Diferentes	Não Diferentes		Não Diferentes			
Média do Grupo 1	597,3333333	7,5	4,5	4,9375	0,0625		
n do Grupo 1	3	16	16	16	16		
Média do Grupo 2	498	3,8125	2,5	2,5	0,5625		
n do Grupo 2	5	16	16	16	16		
Média do Grupo 3	388,25	2,5	1,375	1	0,75		
n do Grupo 3	4	8	8	8	8		
MQD	3868,157407	12,22804054	7,023648649	8,728040541	0,766891892		
Estatística Q	3,95	3,44	3,44	3,44	3,44		
Comparação entre PRINCIPANTES e INTERMEDIÁRIOS							
Diferença Absoluta	99,33333333	3,6875	2	2,4375	0,5		
Erro Padrão da Diferença	32,11710824	0,874215382	0,662554179	0,738581433	0,21893091		
Intervalo Crítico	126,8625776	3,007300913	2,279186377	2,540720131	0,75312233		
Médias dos Grupos 1 e 2 são	Não Diferentes	Diferentes	Não Diferentes	Não Diferentes	Não Diferentes		
Comparação entre PRINCIPANTES e EXPERIENTES							
Diferença Absoluta	209,0833333	5	3,125	3,9375	0,6875		
Erro Padrão da Diferença	33,58887579	1,070690805	0,811459833	0,904573823	0,268134509		
Intervalo Crítico	132,6760594	3,68317637	2,791421826	3,11173395	0,922382711		
Médias dos Grupos 1 e 3 são	Diferentes	Diferentes	Diferentes	Diferentes	Não Diferentes		
Comparação entre INTERMEDIÁRIOS e EXPERIENTES							
Diferença Absoluta	109,75	1,3125	1,125	1,5	0,1875		
Erro Padrão da Diferença	29,5014477	1,070690805	0,811459833	0,904573823	0,268134509		
Intervalo Crítico	116,5307184	3,68317637	2,791421826	3,11173395	0,922382711		
Médias dos Grupos 2 e 3 são	Não Diferentes	Não Diferentes	Não Diferentes	Não Diferentes	Não Diferentes		

Quadro 44 – Síntese gráfica dos resultados do procedimento Tukey-Kramer.

CATEGORIAS DIFERENTES - PROCEDIMENTO DE TUKEY-KRAMER						
Tarefa	Indicadores Objetivos					
	T _e (s)		N _{ai}	N _{oi}	N _{er}	N _{ca}
T0	P	E				
T1			P	I E		
T2			P I	E	P	E
T3			P	E		
T4	P	E	P	I E	P	E

	Principiantes		Intermediários		Experientes
--	---------------	--	----------------	--	-------------

Antes de tudo, vale a pena mencionar que, no tocante aos dados referentes à **Tarefa 0**, o procedimento Tukey-Kramer possibilitou a identificação de uma diferença apenas entre os extremos das categorias de usuários de teste consideradas, i.e. *principiantes* e *experientes*, satisfazendo ao propósito da inclusão desta tarefa no processamento estatístico, conforme explicitado na sub-seção anterior. No tocante aos demais indicadores pré-definidos, não se constatou diferenças entre as diferentes categorias consideradas, conforme previsto.

A Tabela 8 e, sobretudo, a representação sumariada apresentada no Quadro 44 deixam claro que, em essência, não se registrou, em nenhuma das tarefas de teste, diferenças estatisticamente identificáveis entre os indicadores quantitativos relativos às categorias *intermediários* e *experientes*, salvo no número de ações incorretas associado à **Tarefa 2**. Também não se registrou nítidas diferenças entre as categorias *principiantes* e *intermediários*, a não ser aquelas identificadas nas **Tarefas 2 e 4**, concernentes ao número de ações incorretas. Entretanto, todos os testes estatísticos evidenciaram diferenças representativas no desempenho das categorias extremas, i.e. *principiantes* e *experientes*.

7.3.3 Problemas Identificados a partir da Mensuração do Desempenho do Usuário

O processo de observação direta e as análises retrospectivas dos registros em vídeo das sessões de teste do ensaio de usabilidade possibilitaram a identificação de uma série de falhas no processo interativo do usuário com o produto avaliado, algumas das quais também verificadas a partir da inspeção de conformidade do produto às *Partes 14, 15 e 16* do padrão internacional *ISO 9241*. Serão apresentados e discutidos ao longo desta seção apenas os problemas cuja identificação não foi comum aos dois enfoques mencionados. Na próxima sub-seção, que apresentará um sumário dos problemas e um parecer sobre o produto, com base na mensuração do desempenho do usuário, serão destacadas as falhas cuja identificação foi comum aos referidos enfoques.

A primeira falha registrada diz respeito à clareza das mensagens de erro do *MATLAB* v. 5.3.0, estando associada à repetição de erros, pelo menos duas vezes, por **76,2%** dos

participantes das sessões de teste. No início de cada sessão de teste, o participante criou uma pasta de trabalho pessoal em *E:\users\Teste*, para a qual transferiu todos os *scripts* (arquivos *.m*) necessários à realização dos testes (**Tarefa 0**) e na qual armazenou os arquivos de dados criados na **Tarefa 1**. Ao tentar acessar *scripts* ou arquivos de dados de sua pasta de trabalho, e.g. o arquivo *dados1.dat*, sem antes redirecionar o caminho *default* assumido pelo *MATLAB v. 5.3.0*, o usuário deverá receber uma mensagem de erro, uma vez que na pasta *default* não existe nenhum arquivo com o nome *dados1.dat*. No entanto, além de não se destacar visualmente da frase de comando (em cor ou em tipo de fonte), a mensagem oferece pouco ou nenhum auxílio para o usuário, conforme ilustra a Fig. 48.

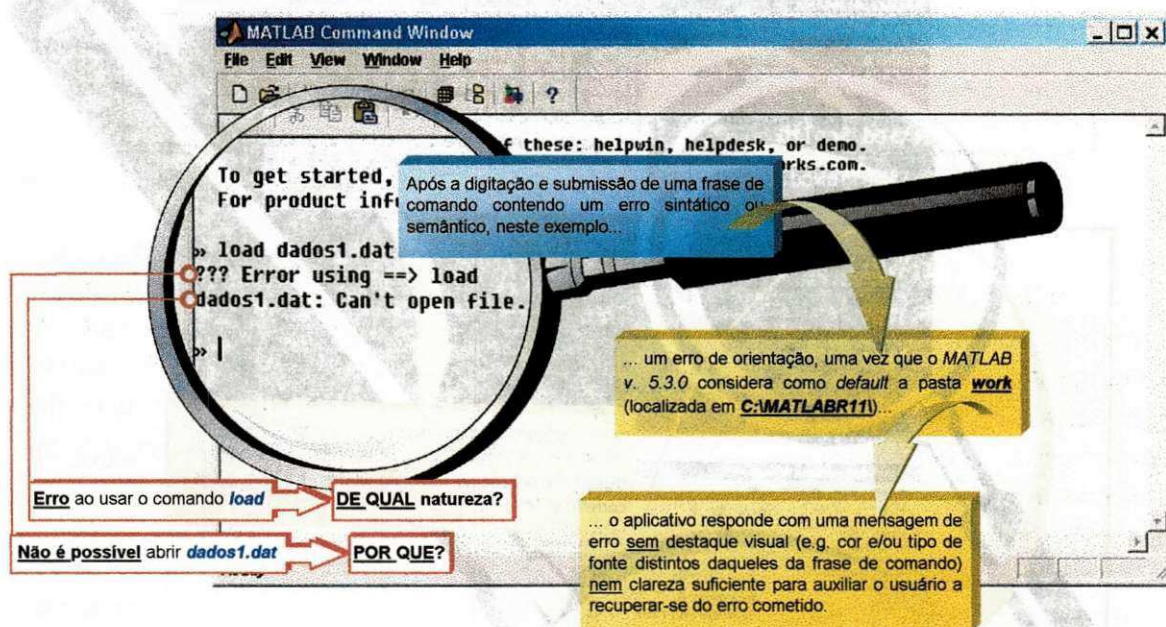


Fig. 48 – Falha do *MATLAB v. 5.3.0* relativa ao destaque e à clareza das mensagens de erros.

Mensagens redigidas neste estilo confundem usuários principiantes e, até mesmo, intermediários e experientes, ao invés de auxiliá-los na recuperação dos erros cometidos. A sugestão para a minimização deste problema é uma revisão da semântica e a adoção de uma estratégia de realce das mensagens de erro, que facilite não somente a compreensão, mas também a discriminação em tela com relação às frases de comando.

A segunda falha, também identificada durante a execução da **Tarefa 0**, está relacionada à semelhança de alguns nomes de comandos do *MATLAB v. 5.3.0* com nomes de comandos do *MS-DOS* (e.g., *dir*, *mkdir*, *cd*, *edit*), assim como dos mecanismos utilizados pelo *MATLAB v. 5.3.0* e pelo *MS-DOS* para executá-los. Tal semelhança, ao mesmo tempo que pode ser útil, pois considera o aprendizado anterior do usuário, também pode ser uma fonte de confusão, nos casos em que há diferenças nas estruturas das linguagens.

Tome-se como exemplo o comando *cd* (*change directory*), empregado para alterar o diretório de trabalho corrente em ambas as aplicações acima referidas. Em ambos os casos, a sintaxe genérica *cd diretório especificado* redireciona o caminho corrente para aquele do diretório especificado. Quando se trata de redirecionar o caminho em nível da raiz da estrutura de

diretórios⁹⁹, o caso do *MS-DOS* executa o redirecionamento apenas a partir do nome da partição (e.g., c:, d:) ou do periférico (a:), dispensando a digitação do nome do comando (Fig. 49).

Tendo em vista que no *MS-DOS* se pode alterar o caminho em nível de raiz da estrutura de diretórios sem ser necessário digitar o nome do comando (*cd*), conforme ilustra a Fig. 47, os usuários ficam confusos ao usar a mesma estratégia na linha de comando do *MATLAB* v. 5.3.0 e recebem uma mensagem de erro.

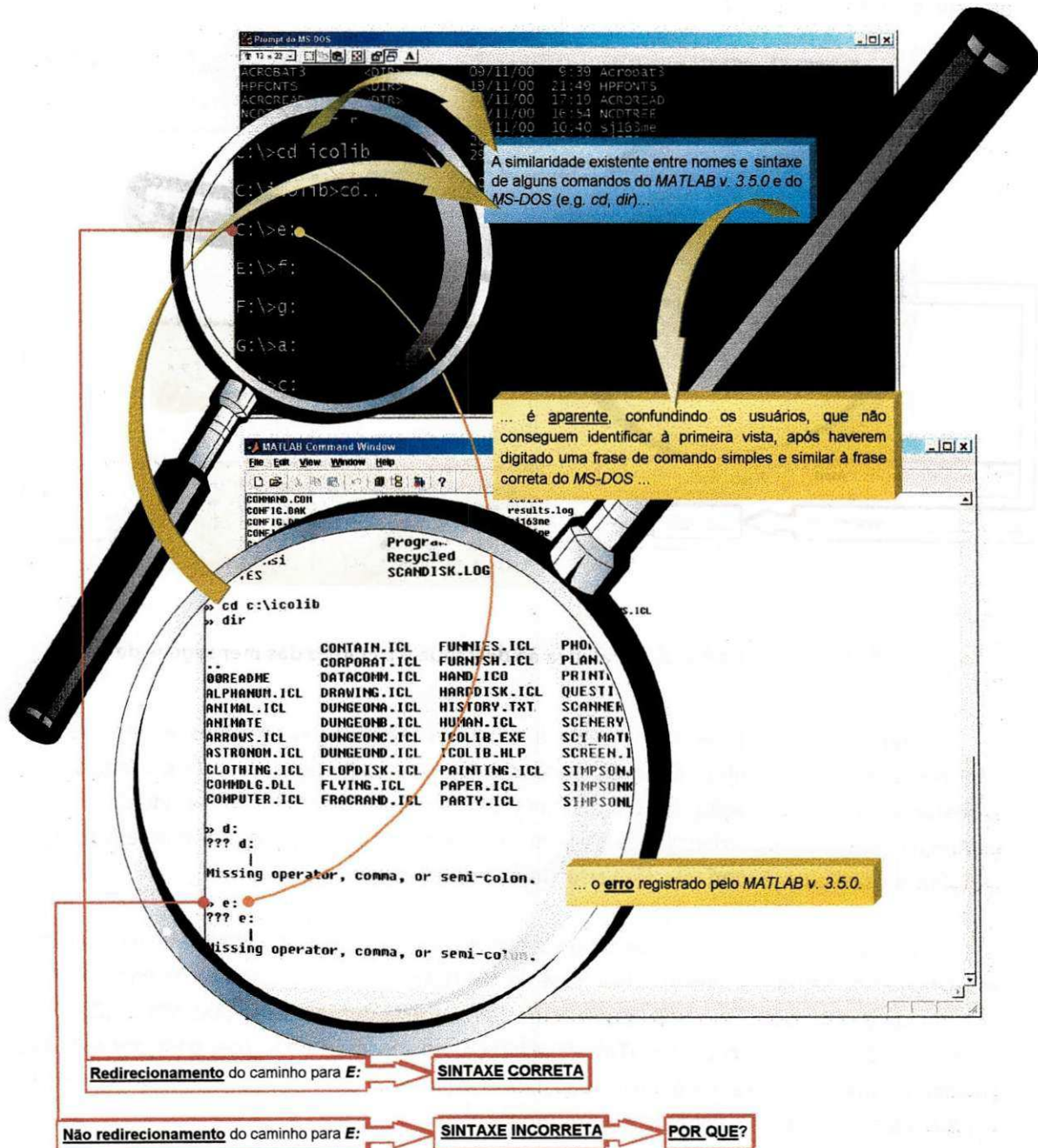


Fig. 49 – Falha do *MATLAB* v. 5.3.0 devido à inconsistência da similaridade de nomes e da sintaxe de frases de comando entre o *MS-DOS* e o *MATLAB* v. 5.3.0.

⁹⁹ E.g., quando o disco rígido possui mais de uma partição, quando se dispõe de mais de um disco rígido no sistema computacional ou quando há diferentes periféricos de armazenamento permanente e temporário instalados.

A falha evidenciada através da Fig. 49 está intimamente relacionada com a falha anterior, como se pode perceber a partir da leitura da mensagem de erro *Missing operator, comma, or semicolon*, que não traduz a natureza real do erro cometido pelo usuário, i.e., o erro de haver empregado uma estrutura sintática correta do ponto de vista do *MS-DOS*, porém não aceita pelo *MATLAB v. 5.3.0*. Esta falha fez com que **83,4%** dos participantes das sessões de teste repetissem o mesmo erro pelo menos duas vezes.

A Fig. 49 também evidencia a terceira falha do produto avaliado, a de não apresentar antes do *prompt* o caminho atual, como o faz o *MS-DOS*. Este lapso obriga o usuário a usar o comando *pwd* (*print work directory*) para mostrar o diretório corrente ou, similarmente ao *MS-DOS*, o comando *dir* (*list directory*) para listar o conteúdo do diretório de trabalho corrente e, indiretamente, descobrir onde realmente se encontra, em nível da estrutura de diretórios e arquivos do sistema, como o fizeram **96,7%** dos usuários de teste.

A quarta falha está associada ao uso do comando *clear* que, conforme visto na sub-seção 7.2.2, é um comando destrutivo que não recebeu a atenção devida da equipe de projeto do *MATLAB v. 5.3.0*. Todavia, a falha detectada a partir da observação direta do usuário de teste diz respeito a outro aspecto do comando *clear*. Quando o usuário executa um *script* com erros ou que utiliza um arquivo de dados que implicam divisões por zero, após a correção dos erros e nova execução do *script*, o *MATLAB v. 5.3.0* continua processando os arquivos como se os erros não houvessem sido corrigidos. Isto resulta da manutenção de resíduos de processamentos anteriores no espaço de trabalho.

Deste modo, torna-se necessária a "limpeza completa" (*clear all*) do espaço de processamento, antes de cada tentativa de reprocessamento de *scripts* e/ou de dados corrigidos, o que nem sempre é óbvio para os usuários, mesmo os experientes. Devido a este problema, durante as sessões de teste registrou-se a repetição do mesmo erro pelo menos duas vezes, para **61,15%** do universo amostral de participantes do experimento.

As falhas que serão descritas a seguir são relativas à janela de figuras do *MATLAB v. 5.3.0*, que utiliza diálogos via menu e manipulação direta. Todas elas estão interrelacionadas, uma vez que dizem respeito à manipulação de atributos das representações visualizadas na janela de gráficos. Observe-se inicialmente a janela apresentada na parte superior esquerda da Fig. 50 e, mais especificamente, as sub-opções da opção *Edit* do menu principal.

Verifica-se que todas sub-opções da opção *Edit* se encontram ativas, exceto *Undo*, *Paste* e *Select All*. Mesmo que o usuário não analise o contexto de trabalho, é óbvio semântica e visualmente que estas últimas não podem ser acessadas, uma vez que estão desativadas.

Observe-se então a janela apresentada na parte inferior direita da Fig. 48, focalizando a atenção na lupa. Todas as opções estão ativas, o que implica dizer que todas são acessíveis, i.e., que se pode acessá-las independentemente de quaisquer outras condições.

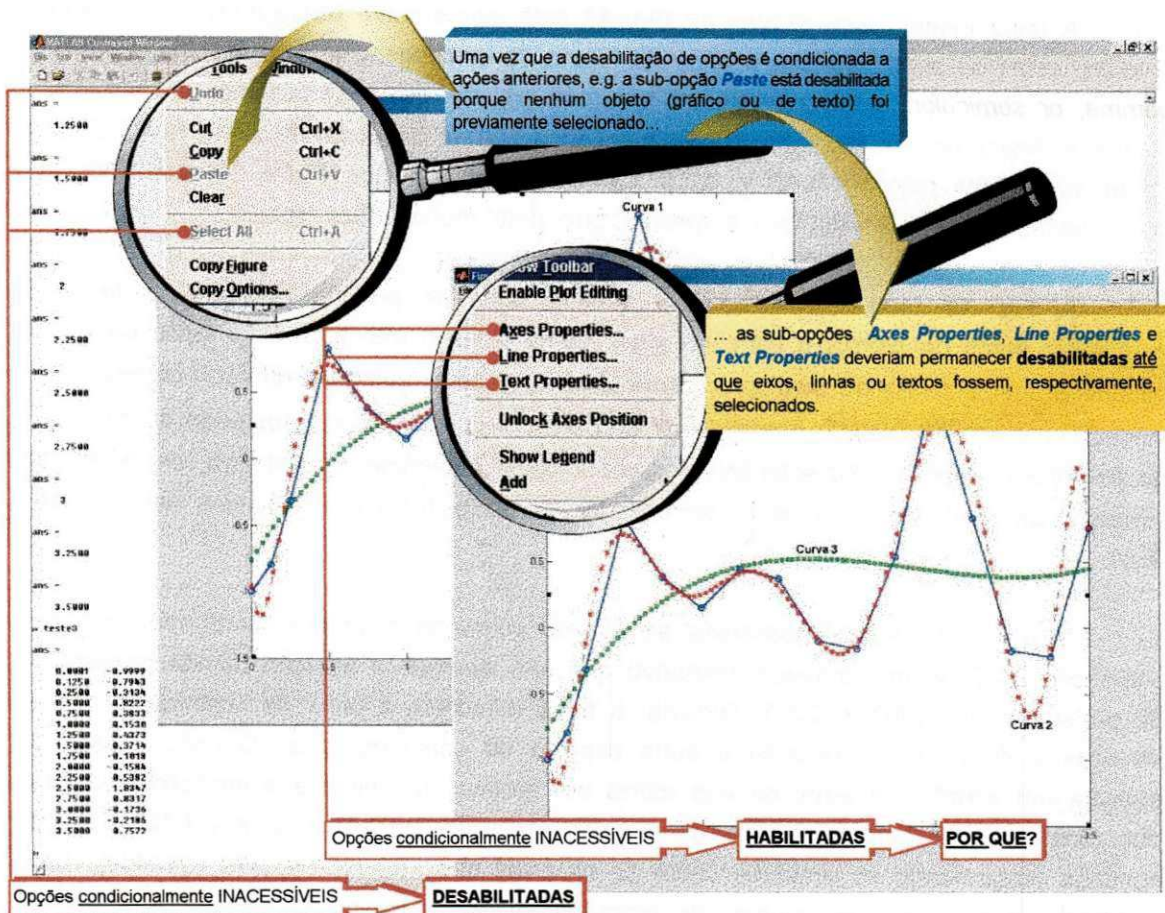


Fig. 50 – Falha do MATLAB v. 5.3.0 relativa à alteração das propriedades de objetos gráficos ou de texto na janela de figuras.

No entanto, se o usuário tentar acessar a opção *Line Properties* ou *Text Properties* sem antes haver selecionado, respectivamente, um objeto gráfico ou um objeto de texto, receberá uma mensagem de erro, informando que nenhum objeto foi previamente selecionado. Esta estratégia de interação confunde o usuário, conforme se pôde perceber nas sessões de teste, nas quais **59,8%** dos participantes tiveram de corrigir, pelo menos duas vezes, a ação incorreta de tentar alterar propriedades de curvas sem antes havê-las selecionado. O mesmo ocorreu com **61,3%** dos participantes, em relação a objetos de texto. Vale a pena ressaltar que o mesmo problema ocorre com a sub-opção *Axes Properties*, cujo acesso implica a seleção anterior dos eixos do gráfico. Sugere-se para o problema uma análise da consistência do projeto do menu.

Associada a esta há outra falha, concernente ao retrocesso a estados ou ações anteriores (*undo*), após a manipulação de alguma propriedade dos eixos, curvas ou textos. Observe-se a *Curva1* nas duas janelas ilustradas na Fig. 51. Na janela superior, a *Curva1* foi traçada em cor azul, uma das cores da paleta *default* do MATLAB v. 5.3.0, enquanto que a janela inferior apresenta a referida curva em amarelo, o que indica que a sub-opção *Line Properties* foi utilizada. Partindo do pressuposto que a sub-opção *Undo* (opção *Edit*) se encontra inicialmente **desabilitada** porque nenhuma ação foi executada, **após** a alteração de uma das propriedades da *Curva1*, no caso a *cor*, a referida sub-opção deveria ser **habilitada**, tornando possível o retrocesso ao estado anterior, i.e. *Curva1* traçada em azul. No entanto, verifica-se que isto **não** ocorre. Assim como para a falha anterior, sugere-se uma análise da consistência do projeto do menu, a fim de que o

problema seja corrigido.

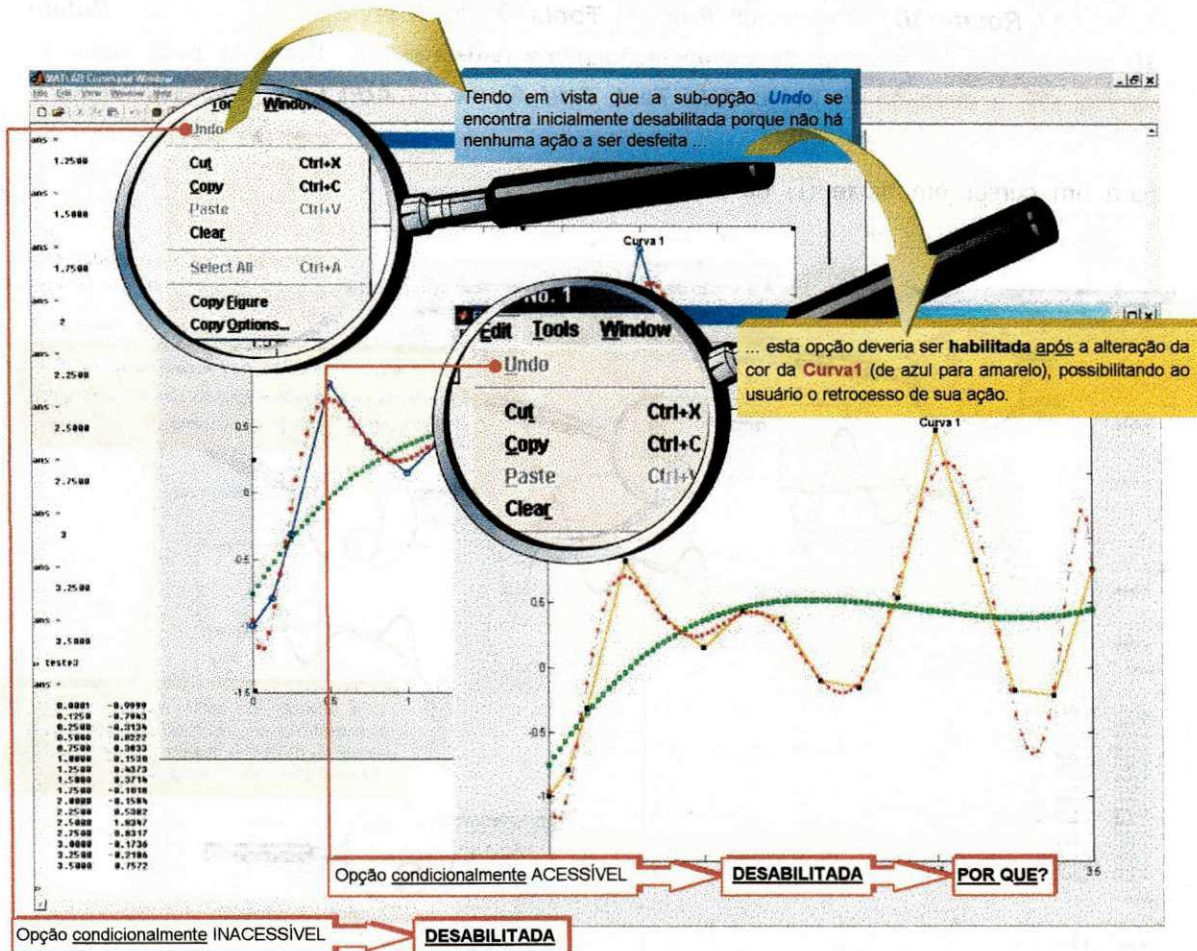


Fig. 51 – Falha do **MATLAB v. 5.3.0** relativa ao retrocesso de ações (*undo*) após a alteração das propriedades de objetos gráficos ou de texto na janela de figuras.

Apesar de não servir como justificativa, poder-se-ia argumentar que a inconsistência explicitada na Fig. 51 seria compensada com a inclusão de um botão de retrocesso de ações na barra de ferramentas de manipulação direta. Todavia, a barra de ferramentas de manipulação direta da janela de figuras do **MATLAB v. 5.3.0** não inclui tal botão, conforme ilustrado na Fig. 52.

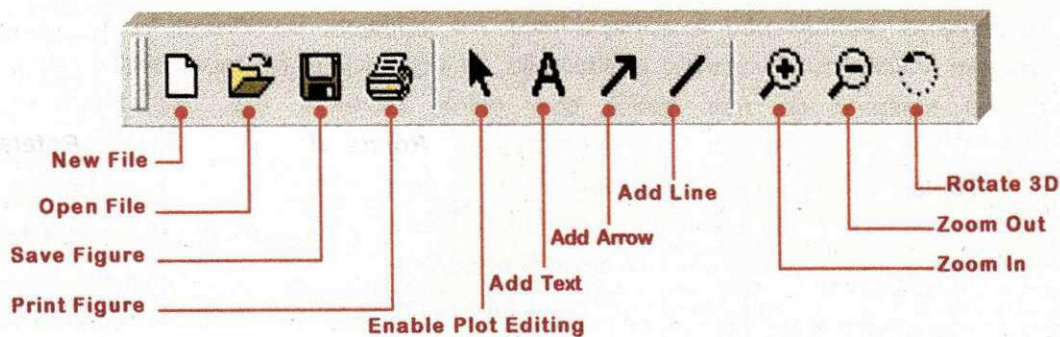


Fig. 52 – Barra de ferramentas de manipulação direta da janela de figuras do **MATLAB v. 5.3.0**.

A oitava falha identificada diz respeito à rotação tridimensional do sistema de eixos das representações gráficas exibidas na janela de figuras do *MATLAB* v. 5.3.0, a partir da sub-opção **Rotate 3D**, subordinada à opção **Tools** do menu principal, ou do botão **Rotate 3D** da barra de ferramentas de manipulação direta (vide Fig. 52). Como se pode observar na Fig. 53, a ativação do botão **Add Text**, **Add Arrow** ou **Add Line** é seguida de uma indicação visual para o usuário – a transformação do cursor em seta (↖) respectivamente para um cursor em haste (I) ou em cruz (+).

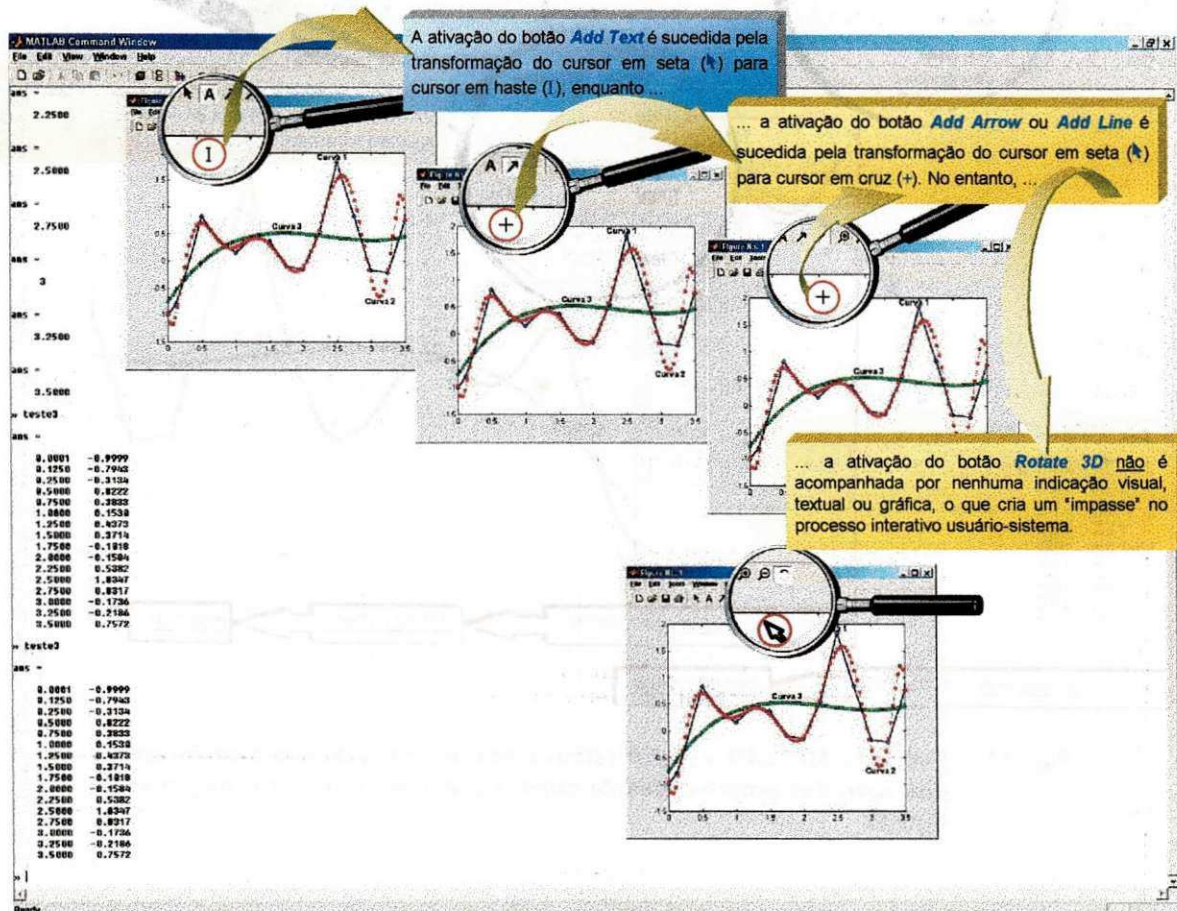


Fig. 53 – Falha do *MATLAB* v. 5.3.0 relativa indicação visual da ativação do recurso de rotação tridimensional do sistema de eixos da janela de figuras.

Entretanto, a ativação do botão de rotação tridimensional do sistema de eixos das representações gráficas não é acompanhada por nenhuma transformação do cursor ou mensagem solicitando ao usuário a alteração da perspectiva via manipulação direta dos eixos, ao contrário do que ocorre com outros produtos, e.g., *MS Word2000* ou *CorelDRAW 9*. Deste modo, após a ativação do botão **Rotate 3D** ou da sub-opção **Rotate 3D** (opção **Tools** do menu principal), o aplicativo fica aguardando o ajuste da perspectiva pelo usuário, a partir do posicionamento do cursor sobre a área do gráfico, seguido do arrasto dos eixos até que a *elevação* e o *azimute* do plano do gráfico estejam conforme o interesse do usuário. Este, por sua vez, permanece aguardando que o aplicativo forneça uma indicação visual de que está pronto para executar a rotação dos eixos do gráfico. Tal situação pode criar um “impasse” no processo interativo, o que ocorreu com **89,7%** dos participantes do ensaio de usabilidade, para os quais foi necessária a intervenção do avaliador.

A Fig. 54 ilustra a discriminação visual adotada pelo *MS Word2000* e pelo *CorelDRAW 9* para as ações de redimensionamento e rotação de objetos gráficos. Do lado esquerdo da Fig. 54, observa-se que o *MS Word2000* apresenta as alças de redimensionamento (□) sempre que um objeto é criado ou selecionado. Além disso, fora da área do objeto o cursor assume a forma de haste (I), alterando-se para indicar a possibilidade (i) de arrasto (✚), se posicionado no interior da área do objeto, ou (ii) de redimensionamento unidimensional (↔ ou ⇅) ou bidimensional (↔↕), se posicionado sobre as alças de redimensionamento. A rotação do objeto torna-se possível a partir da ativação do botão **Free Rotate** da barra de ferramentas **Drawing**, após o que as arestas do objeto são realçadas por círculos preenchidos em verde (●). Além disso, fora da área do objeto o cursor assume uma forma que sugere rotação (⤵), alterando-se para indicar a possibilidade (i) de arrasto (✚), se posicionado no interior da área do objeto.

Quanto ao *CorelDRAW 9*, este aplicativo apresenta as alças de redimensionamento (■) sempre que um objeto é criado ou selecionado. Embora não mude de forma dentro e fora da área do objeto, o cursor é alterado para indicar a possibilidade de redimensionamento unidimensional (↔ ou ⇅) ou bidimensional (↔↕), se posicionado sobre as alças de redimensionamento. A rotação do objeto torna-se possível a partir de um clique duplo do botão esquerdo do *mouse* dentro da área do objeto, após o que as arestas do objeto são realçadas por segmentos com setas em ambas as extremidades (↶, ↷, ↕ e ↔), os quais sugerem a rotação nos sentidos horário e anti-horário, conforme ilustrado na porção direita da Fig. 54.

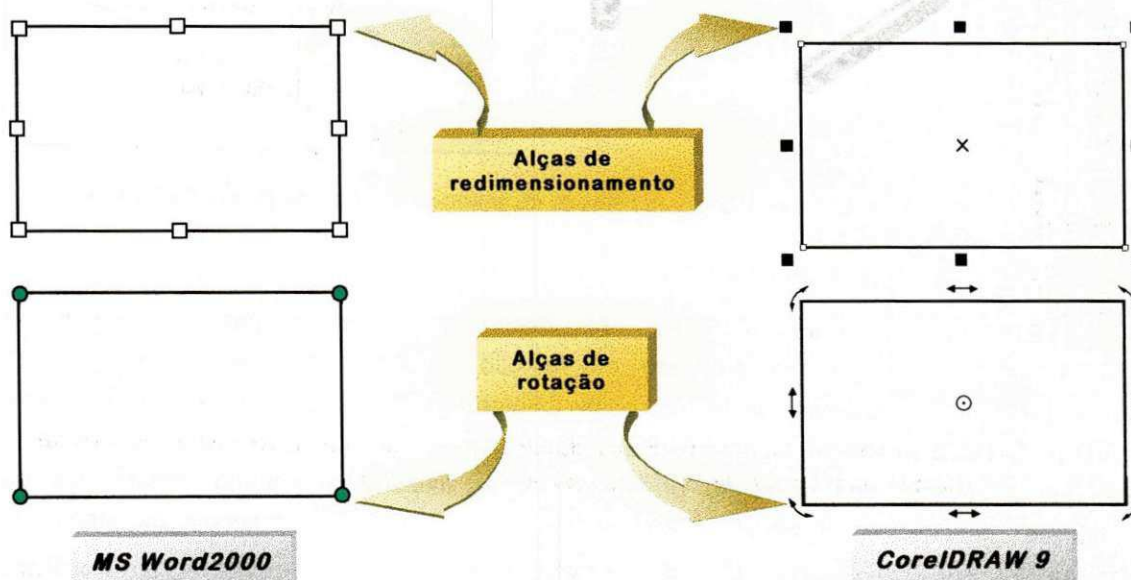


Fig. 54 – Estratégias de indicação visual nos processos de redimensionamento e rotação tridimensional de objetos no (a) *MS Word2000* e (b) *CorelDRAW 9*.

Diante do exposto, sugere-se um reprojeto do mecanismo de rotação tridimensional da interface da janela de figuras do *MATLAB v. 5.3.0*, visando minimizar os efeitos do mecanismo atual sobre o processo interativo usuário-aplicativo e, por conseguinte, evitar o "impasse" anteriormente mencionado.

A nona falha é relativa ao mecanismo pouco eficiente de ajuste do *azimute* e da *elevação* do plano do gráfico da janela de figuras do *MATLAB v. 5.3.0*, que será

detalhada a partir da Fig. 55.

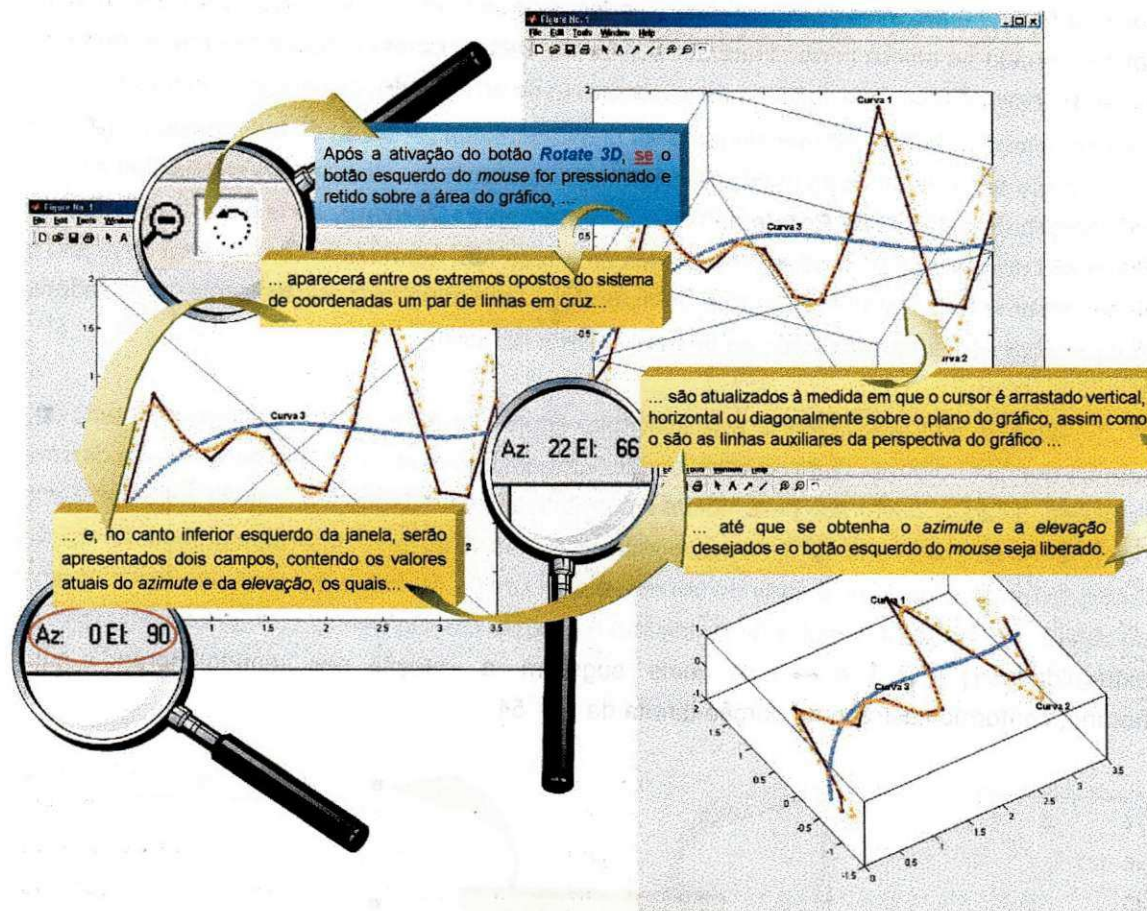


Fig. 55 – Detalhamento do processo de rotação do plano do gráfico da janela de figuras do *MATLAB v. 5.3.0*.

Embora as vinhetas da Fig. 55 descrevam passo a passo o ajuste do *azimute* e da *elevação* no processo de rotação do plano do gráfico, vale a pena comentar alguns aspectos deste processo. Conforme comentado no registro da falha anterior, após permanecer alguns instantes aguardando uma indicação visual do aplicativo de que a rotação dos eixos do gráfico pode ser iniciada, o usuário tentará sair do “impasse” consultando a ajuda ou simplesmente explorando os recursos da janela de figuras.

Observou-se em **63,9%** dos participantes do ensaio que a primeira decisão tomada foi pressionar o botão esquerdo do *mouse*, após posicionar o cursor sobre a área do gráfico. No entanto, como não foi fornecida pelo aplicativo nenhuma indicação visual além do par de linhas auxiliares traçado entre os extremos opostos do sistema de eixos, **51,2%** do referido percentual de usuários de teste liberaram o botão esquerdo do *mouse*, ao invés de retê-lo e arrastar o plano do gráfico até a perspectiva desejada.

Os campos contendo os valores do *azimute* e da *elevação* passaram despercebidos à primeira vista por **71,8%** dos usuários de teste, tanto devido à posição em que foram apresentados na tela, quanto pela insuficiência de destaque visual (tipo, tamanho e/ou estilo) da fonte

empregada. Entretanto, o cerne do problema residiu no ajuste dos valores do *azimute* e da *elevação* solicitados no roteiro da **Tarefa 4**. Uma vez que o deslocamento retilíneo do *mouse* na horizontal ou na vertical, sobre o plano do gráfico, é uma atividade que exige uma boa dose de perícia do usuário, registrou-se durante as sessões de teste uma tendência ao desajuste do valor do *azimute* durante as tentativas de ajustar a *elevação* e vice-versa.

Quanto ao ajuste simultâneo deste par de atributos, **94,6%** dos participantes classificaram como crítico o mecanismo de rotação adotado na janela de figuras do *MATLAB v. 5.3.0*. Para a otimização deste mecanismo, sugere-se a desvinculação do ajuste, quer a partir do fornecimento de dois campos para a digitação dos valores desejados ou da inclusão de recursos de manipulação direta para o ajuste por passo (como ocorre no *MS Word2000*) ou contínuo, reformulando a própria estratégia adotada na janela de figuras do *MATLAB v. 5.3.0*.

A décima e última falha não identificada a partir da inspeção de conformidade do produtos ao padrão internacional *ISO 9241*, porém observada durante o ensaio de usabilidade, foi verificada conjuntamente à falha detalhada no parágrafo anterior. Apesar de intimamente relacionada com a sexta e sétima falhas (retrocesso a estados anteriores), esta falha só foi observada durante o ajuste do *azimute* e da *elevação*, razão pela qual preferiu-se registrá-la por último.

A falha foi identificada na trigésima segunda sessão de teste, durante a execução da **Tarefa 4** por um usuário *experiente*. Até então, nenhum usuário de teste tentara desfazer os ajustes do *azimute* e da *elevação*, após o lapso de liberar o botão esquerdo do *mouse* antes de obter a perspectiva desejada. Apesar da experiência de uso do *MATLAB* ao longo de suas versões, o usuário ainda não percebera que a barra de ferramentas de manipulação direta da janela de figuras da versão *5.3.0* não continha um botão para o retrocesso a estados anteriores (*undo*), conforme explicitado anteriormente no registro da sétima falha. Ao acessar a opção **Edit** do menu principal, na tentativa de desfazer o lapso, o usuário verificou que, apesar de habilitada, a sub-opção **Undo** não lhe permitia retroceder ao estado anterior.

O mais grave é que ao tentar sucessivas solicitações de retrocesso, o usuário não obteve nenhuma resposta do aplicativo, ou seja, a sub-opção **Undo** se mostrou supérflua e inútil ao contexto em que a falha foi identificada. Após a referida sessão de teste, pôde-se constatar, a partir de uma análise mais detalhada do problema, que a falha existe desde a abertura da janela de figuras.

Seja através de uma instrução de traçado de gráficos contida em um *script* do *MATLAB*, como foi o caso, seja mediante a solicitação de criação de um novo arquivo de figura ou de abertura de um arquivo de figura já existente, a janela de figuras é aberta com todas as sub-opções da opção **Edit** habilitadas, conforme atesta a Fig. 56. A lupa posicionada sobre as sub-opções da opção **Edit** na janela posterior, a primeira da seqüência, focaliza as sub-opções habilitadas, enquanto a lupa sobre a janela anterior evidencia a desabilitação da sub-opção **Undo** após a inserção de um novo objeto gráfico no interior da área do gráfico, a linha traçada em azul, que aparece parcialmente oculta

pelo painel de menu *Edit*.

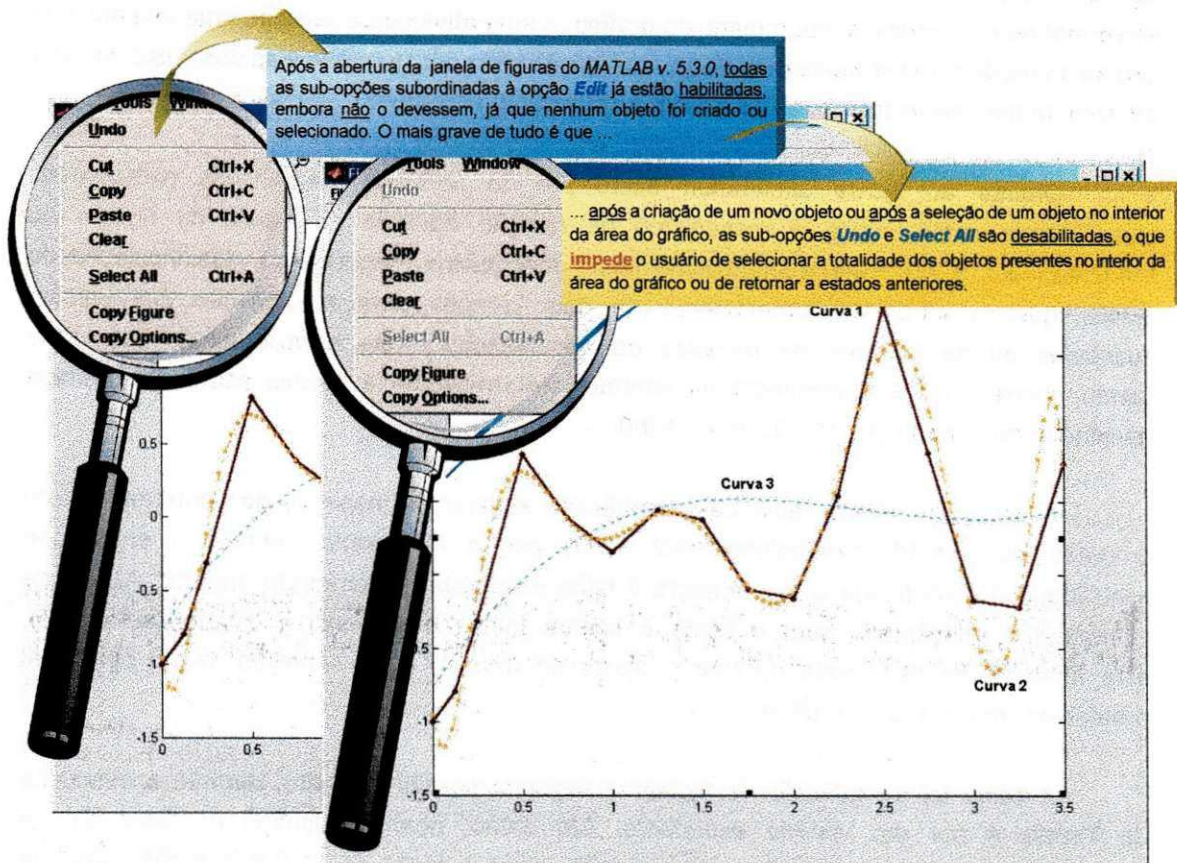


Fig. 56 – Destaque para a falha de projeto do painel de sub-opções da opção *Edit*.

Deve-se, pois, observar que a sexta falha identificada a partir deste enfoque avaliatório é, na realidade, uma extensão desta última falha, i.e. ambas as falhas refletem um projeto inadequado e inconsistente do painel de menu da opção *Edit*. Tendo em vista que traduzem facetas diferentes de uma falha maior, preferiu-se relatá-las separadamente, assim como o foi sua identificação. Todavia, figurarão com uma única falha no sumário de falhas que será apresentado na sub-seção seguinte.

7.3.4 Parecer sobre o Produto com base na Mensuração do Desempenho do Usuário

É conveniente relembrar que, além das falhas relatadas nos parágrafos anteriores, o processo de observação direta e as análises retrospectivas dos registros em vídeo das sessões de teste do ensaio de usabilidade também permitiram identificar uma parte das falhas anteriormente relatadas na seção referente à inspeção de conformidade do produto às *Partes 14, 15 e 16* do padrão internacional ISO 9241.

Uma nova observação do Quadro 14 e, mais especificamente, das informações contidas na coluna *Classe de Problemas Encontrados* (Capítulo 4, seção 4.3), mostra que Queiroz e Tumell [Quei96] adotaram uma classificação em dois níveis para problemas identificados a partir de diferentes técnicas de avaliação.

Um dos níveis da referida classificação diz respeito aos efeitos do problema sobre as ações do usuário, através da qual os problemas identificados são rotulados como *superficiais*, *intermediários* ou *graves*, conforme o grau de desconforto que causam ao usuário e o comprometimento do curso das ações que estes executam. Adicionalmente, o outro nível da classificação, a partir do qual os problemas são rotulados como *de consistência*, *recorrentes* e *gerais*, é concernente aos efeitos de cada problema identificado sobre o funcionamento do sistema e sobre o acesso a suas facilidades pelo usuário.

O Quadro 45 sintetiza os dois níveis da classificação mencionados, inicialmente com base em informações adaptadas de compilações feitas por Jeffries *et al.* [Jeff91], Nielsen [Niel93c] e Van Vianen *et al.* [VanV96] sobre técnicas de avaliação. Vale salientar que estas informações foram posteriormente revisadas e complementadas com outras mais recentes, compiladas por Dix *et al.* [Dix98], Jordan [Jord98] e Constantine e Lockwood [Cons99].

Quadro 45 – Síntese da classificação adotada por Queiroz e Turnell [Quei96] para problemas identificados em interfaces homem-máquina.

NÍVEL	CLASSE DO PROBLEMA	DESCRIÇÃO
1	Superficial	Causa desconforto ao usuário, porém não compromete a execução das ações intentadas, exigindo-lhe apenas um processo de adaptação. Tal fato não implica que o problema de usabilidade não deva ser solucionado, partindo do pressuposto que não é o usuário quem deve se "amoldar" às características do produto e, sim, o inverso (e.g., falta de clareza em mensagens de erro, ausência de realimentação no processo interativo).
	Intermediário	Causa desconforto ao usuário, além de forçá-lo a alterar o curso de suas ações para atingir o propósito almejado, e.g., uma falha no acesso de uma opção ou sub-opção do menu de uma interface multimodal (menus, linguagem de comandos e manipulação direta) destinados a diferentes categorias de usuários [Quei94].
	Grave	Causa grande desconforto ao usuário, por comprometer seriamente a execução de um curso de ações, e.g., o travamento de uma função que exija o reinício do processo (travamento <i>parcial</i>) ou de todo o sistema (travamento <i>total</i>).
2	De consistência	Relativo a conflitos entre partes do sistema avaliado, tanto em nível estrutural e estético, quanto semântico e operacional, e.g., apresentação de mensagens de erro referentes à mesma função do aplicativo ou funções afins em diferentes regiões da tela em cores distintas (estrutura e estética) e/ou apresentação de mensagens de erro semanticamente divergentes relativas à mesma função do aplicativo ou funções afins (semântica).
	Recorrente	Interfere no processo interativo a cada vez que se repetem determinadas condições operacionais, e.g., a "quebra" do retorno de informações ao usuário cada vez que este solicita a aquisição de parâmetros estatísticos relativos a uma imagem analisada em um sistema de processamento de imagens.
	Geral	Afeta várias partes do sistema, e.g., falhas que induzem panes parciais ou totais no sistema.

Como se pôde observar ao longo da seção 7.2 e da sub-seção anterior, as falhas identificadas, tanto através da inspeção de conformidade às *Partes 14, 15 e 16* do padrão *ISO 9241* quanto da mensuração do desempenho do usuário, se enquadram nas classes *superficial* ou *intermediário*, de acordo com o primeiro nível de classificação, assim como podem ser rotuladas como *de consistência* ou *recorrentes*, conforme o segundo nível da referida classificação. O Quadro 46 reforça esta afirmativa, listando de forma resumida os problemas identificados a partir da mensuração do desempenho do usuário.

Quadro 46 – Sumário de Falhas e Parecer sobre o **MATLAB v. 5.3.0** com base na **Mensuração do Desempenho do Usuário**.

# DA FALHA	COMENTÁRIO
01	Não há distinção entre objetos <i>não selecionáveis (títulos e textos explicativos)</i> e objetos <i>selecionáveis (itens consultáveis)</i> nas modalidades de ajuda <i>Help Windows</i> e <i>Help Tips</i> .
02	Várias palavras de comando <u>excedem</u> o limite de 7 caracteres estabelecido no padrão (há palavras com até 19 caracteres).
03	Não há nenhum mecanismo que torne óbvio para o usuário o uso de <i>teclas de funções</i> e de <i>atalhos</i> no diálogo via linha de comando.
04	Não é permitida a redigitação ou edição <u>apenas</u> de trechos incorretos de comandos e parâmetros associados.
05	O sistema <u>não</u> interpreta nem aceita a entrada de comandos contendo erros de digitação e/ou ortografia, mesmo que não haja ambigüidade quanto ao comando pretendido.
06	Não há mecanismos para o processamento de comandos destrutivos (e.g., <i>delete</i> , <i>clear</i>).
07	Não há mecanismos para a reversão de ações do usuário (<i>undo</i>).
08	O sistema executa <u>qualquer</u> sentença de comando correta <u>sem</u> solicitar antes nenhuma confirmação do usuário.
09	O sistema <u>não</u> oferece nenhuma realimentação sobre o tempo restante para a conclusão de processos em andamento.
10	O sistema <u>não</u> evidencia os erros cometidos em linhas de comando (suas mensagens de erro são lacônicas e pouco esclarecedoras).
11	O sistema <u>não</u> explicita a natureza (e.g. semântica, sintática, de orientação) do erro cometido, nem destaca <u>visualmente</u> as mensagens de erro das frases de comando, o que, em diversas situações, desnorreia o usuário e dificulta sua recuperação face aos erros cometidos.
12	A linguagem de comandos utiliza alguns nomes e sintaxes de comandos similares àquelas utilizadas pelo <i>MS-DOS</i> , porém <u>não mantém a consistência</u> em todas as circunstâncias em que se constata tal similaridade, o que confunde o usuário e propicia o cometimento de erros.
13	O <i>prompt</i> da linha de comando <u>não</u> apresenta o caminho corrente, forçando o usuário a usar os comandos <i>pwd</i> e <i>dir</i> para "se situar" na estrutura de diretórios (dispêndio de tempo extra).
14	Não há mecanismo de acesso à barra de ferramentas de manipulação direta via teclado (dispositivo de entrada alternativo do <i>mouse</i>)
15	Os botões <i>Home</i> e <i>Tips</i> nas janelas de ajuda <i>Help Windows</i> e <i>Help Tips</i> <u>contrariam</u> a seqüência de manipulação direta de objetos, estando habilitados em circunstâncias em que não o deveriam.
16	A barra de ferramentas de manipulações diretas do <i>Editor/Debugger</i> <u>não possui</u> ícone de retrocesso de ações (<i>undo</i>), o que impede o retorno do usuário a estados anteriores.
17	Antes de cada tentativa de reproprocessamento de <i>scripts</i> e/ou de dados corrigidos, é necessário o uso do comando <i>clear</i> para a eliminação de resíduos de dados de processamentos anteriores no espaço de trabalho. Este requisito <u>não é óbvio</u> , forçando usualmente o usuário a "quebrar" a seqüência <u>direta</u> de ações, o que propicia o cometimento de erros de processamento de dados e pode confundir o usuário.
18	O projeto do painel de menu da opção <i>Edit</i> da janela de figuras é <u>inconsistente</u> , apresentando sub-opções em estados (habilitadas ou desabilitadas) que <u>contrariam</u> tanto os contextos de uso quanto os estados das sub-opções de outros painéis do menu principal da janela.
19	O retrocesso de ações (<i>undo</i>) executadas sobre objetos localizados na janela de gráficos, a partir da sub-opção <i>Undo</i> do menu (painel da opção <i>Edit</i>) da janela de figuras, <u>não é possível</u> , uma vez que a sub-opção permanece desabilitada independentemente das ações executadas na área do gráfico (construção de linhas, alteração de propriedades dos eixos, de objetos gráficos ou de objetos de texto), o que usualmente "quebra" a seqüência <u>direta</u> de ações do usuário.
20	A barra de ferramentas de manipulações diretas da janela de figuras <u>não possui</u> ícone de retrocesso de ações (<i>undo</i>), o que impede o retorno do usuário a estados anteriores.
21	Não é oferecida uma indicação visual <u>óbvia</u> das ações de manipulação direta para a rotação tridimensional do plano do gráfico consecutivas à ativação do botão <i>Rotate 3D</i> da barra de ferramentas (ou da sub-opção de mesmo nome do painel de menu da opção <i>Tools</i>).
22	A rotação tridimensional é <u>crítica</u> , i.e. o mecanismo de ajuste dos atributos <i>azimute</i> e <i>elevação</i> é <u>pouco eficiente</u> , implicando tempo e esforço extras dos usuários.
PARECER: Similarmente à inspeção de conformidade do MATLAB v. 5.3.0 , a mensuração do desempenho do usuário através do ensaio de usabilidade demonstrou que apesar dos recursos poderosos para a solução de diversos contextos de problemas, o produto apresenta uma série de falhas na interface (acima listadas), as quais merecem ser convenientemente analisadas e corrigidas, a fim de que o processo de interação usuário-produto seja otimizado.	

Ao final do Quadro 46, similarmente ao Quadro 40, fornece-se um parecer que reflete o ponto de vista deste enfoque avaliatório. Antes de encerrar a seção, três comentários se fazem necessários. O primeiro deles diz respeito à informação contida na célula da coluna *Classe de Problemas Encontrados* do Quadro 14 referente à técnica de avaliação *Observação*. Segundo esta informação, a técnica se presta à identificação de falhas *graves gerais* e *recorrentes* na interface. Obviamente, esta afirmação será verdadeira se estas classes de falhas existirem. Caso contrário, a técnica poderá ser empregada, de modo igualmente efetivo, para a identificação de outras classes de falhas, conforme demonstraram os resultados apresentados nesta seção.

O segundo comentário se refere à descrição da classe *superficial* do Quadro 45 e sua aplicação ao contexto da avaliação do *MATLAB v. 5.3.0*, a partir da mensuração do desempenho do usuário e da inspeção de conformidade deste produto às *Partes 14, 15 e 16* do padrão *ISO 9241*. Embora não tenha sido constatada nenhuma falha *geral grave*, as falhas *superficiais* identificadas não devem ser negligenciadas. É necessário solucioná-las, a fim de que a interface e, por extensão, o produto sejam otimizados, um dos principais requisitos para sua aceitação pela comunidade usuária, preferivelmente aos produtos concorrentes.

O terceiro e último comentário diz respeito um aspecto de grande importância para a aceitação de produtos internacionais: o projeto da interface do produto considerando questões lingüísticas e culturais do país onde se pretende sua difusão. Durante as sessões de teste, constatou-se que **62,1%** dos participantes apresentou, pelo menos uma vez, dificuldades na compreensão dos termos em inglês. Embora este aspecto não constitua diretamente uma falha do produto, a questão da internacionalização é um aspecto que não pode ser relegado ao segundo plano. DeGaldo [DeIG96] e Nielsen [Niel97, Niel00], dentre diversos outros pesquisadores, alertam os desenvolvedores para a questão da internacionalização, à qual é fortemente condicionada a aceitação de produtos e, mais especificamente, de produtos de *software*.

7.4 Resultados da Sondagem da Opinião do Usuário

Esta seção será dividida em duas sub-seções. A primeira delas tratará dos resultados obtidos com a aplicação do *DePerUSI*, o instrumento de sondagem do perfil do usuário, desenvolvido e utilizado antes da condução do ensaio de usabilidade, cuja concepção foi descrita no Capítulo 6 e cujos resultados foram apresentados e discutidos na seção anterior. A partir dos resultados da aplicação do *DePerUSI*, será esboçado o perfil do usuário típico do universo amostral considerado nesta pesquisa, relevante tanto para a contextualização do ensaio de usabilidade quanto para as conclusões que advirão do processamento dos dados coletados. A segunda sub-seção, por sua vez, tratará dos resultados obtidos com a aplicação do *OpUS*, o instrumento de sondagem concebido na fase de elaboração do ensaio de usabilidade, conforme descrito no Capítulo 6, e aplicado imediatamente após os testes de mensuração do desempenho do usuário durante o uso do produto avaliado.

7.4.1 Resultados da Análise do *DePerUSI*

O Quadro 47 sumaria a totalização das respostas dos usuários de teste aos itens do instrumento de delineamento do perfil do usuário de sistemas interativos, reproduzindo sua estrutura, a fim de facilitar a associação dos itens às cifras computadas.



ITEM	OPÇÕES							
Você é:	PÓS-GRADUADO	11	GRADUADO	5	ESTUDANTE DE GRADUAÇÃO	16	ESTUDANTE DE PÓS-GRADUAÇÃO	8
Você é do sexo:	MASCULINO	34	FEMININO	6				
Você é:	DESTRO	36	CANHOTO	3	AMBIDESTRO	1		
Você usa óculos ou lentes de contato?	SIM	32	NÃO	8				
Você pertence à faixa etária de:	18 A 24 ANOS	18	25 A 34 ANOS	14	35 A 45 ANOS	6	ACIMA DE 45 ANOS	2
A natureza da principal atividade que v. desenvolve com o auxílio do PRODUTO é essencialmente de:	PESQUISA E DESENVOLVIMENTO	17	APLICAÇÃO IMEDIATA DE UTILITÁRIOS	0	TREINAMENTO E APRENDIZAGEM	23		
No contexto de suas atividades v. utiliza o PRODUTO em:	Construção de Modelos	9	Verificação de Modelos Existentes	7	Utilização em Teses/Dissertações	12	Uso em Trabalhos de Iniciação Científica	4
	Prestação de Serviços	0	Execução de Projetos	0	Desenvolvimento de Produtos	0	Treinamento em Nível de Extensão	0
	Suporte a Disciplinas Teóricas (Graduação)	15	Suporte a Disciplinas Teóricas (Pós-graduação)	9	Suporte Laboratorial (Graduação)	16	Suporte Laboratorial (Pós-graduação)	5
Sua plataforma computacional é:	UNIX	0	PC	40	MAC	0	OUTRA	0
Você tem experiência prévia com sistemas computacionais?	SIM	40	NÃO	0				
Há quanto tempo v. usa sistemas computacionais?	MENOS DE 3 MESES	0	DE 3 MESES A 1 ANO	1	MAIS DE 1 ANO	39		
Com que frequência v. usa sistemas computacionais?	DIARIAMENTE	33	MENOS DE 1 VEZ POR SEMANA	3	1 VEZ POR SEMANA	4	MENOS DE 1 VEZ POR SEMANA	0
	1 VEZ POR QUINZENA	0	MENOS DE 1 VEZ POR QUINZENA	0	1 VEZ POR MÊS	0	É A 1ª VEZ QUE USO O PRODUTO	0
Você tem experiência prévia com o PRODUTO?	SIM	40	NÃO	0				
Há quanto tempo v. usa o PRODUTO?	MENOS DE 3 MESES	11	DE 3 MESES A 1 ANO	7	MAIS DE 1 ANO	22		
Com que frequência v. usa o PRODUTO?	DIARIAMENTE	6	MENOS DE 1 VEZ POR SEMANA	12	1 VEZ POR SEMANA	4	MENOS DE 1 VEZ POR SEMANA	1
	1 VEZ POR QUINZENA	5	MENOS DE 1 VEZ POR QUINZENA	3	1 VEZ POR MÊS	9	É A 1ª VEZ QUE USO O PRODUTO	0
Qual a versão do PRODUTO que v. utiliza atualmente?	v. 5.3	12	v. 5.2	21	v. 5.1	4	v. 5.0	1
	v. 4x	2	v. 3x	0	v. 2x	0	NUNCA USEI NENHUMA VERSÃO	0
Qual a forma de ajuda do PRODUTO que v. costuma utilizar mais frequentemente?	MANUAL	6	AJUDA DE LINHA DE COMANDO	22	HELPWIN	4	HELPDESK (CD-ROM)	0
	HELPDESK (INTERNET)	1	DEMOS	0	CONSULTAS ORAIS A OUTROS USUÁRIOS	7	SUORTE (FONE-EMAIL) TECNICO	0
Quais as Caixas de Ferramentas (Toolboxes) que v. costuma utilizar mais frequentemente?	Chemometrics	7	Communications	7	Control System	12	Frequency Domain System Identification	0
	Fuzzy Logic	0	Higher-Order Spectral Analysis	0	Image Processing	3	LMI Control	0
	Mapping	0	Predictive Model Control	1	NAG Foundation	0	Neural Network	1
	Optimization	0	Partial Differential Equation	0	OFT Control Design	0	Robust Control	0
	RTW	0	Signal Processing	11	Spline	0	Statistics	1
	Symbolic/Extended Math	0	System Identification	4	Wavelet	3	μ-Analysis and Synthesis	0
Você tem experiência prévia com produtos similares?	SIM	19	NÃO	21				
Se v. respondeu NÃO ao item anterior, encerre neste ponto o preenchimento do questionário. Caso contrário, especifique o produto.		13	MATHEMATICA					
		4	MAPLE					
		1	SCILAB					
		1	CENSTAT					
Se v. preencheu o item anterior, especifique o tempo de uso do produto.	MENOS DE 3 MESES	8	DE 3 MESES A 1 ANO	7	MAIS DE 1 ANO	4		
Se v. tem experiência com um segundo produto similar, especifique-o.		1	MAPLE					
Se v. preencheu o item anterior, especifique o tempo de uso do produto.	MENOS DE 3 MESES	0	DE 3 MESES A 1 ANO	0	MAIS DE 1 ANO	1		
Se v. tem experiência com um terceiro produto similar, especifique-o.		1	PTOLEMY CLASSIC					
Se v. preencheu o item anterior, especifique o tempo de uso do produto.	MENOS DE 3 MESES	0	DE 3 MESES A 1 ANO	0	MAIS DE 1 ANO	1		



O Produto - Uso e Navegação											
ASPECTO	ESCALA					NR	ESCORE BRUTO	ESCORE ESPERADOS	ESCORE BRUTO OBTIDO (Seção 1)	ESCORE PONDERADO OBTIDO (Seção 1)	
	1	2	3	4	5						
	Muito fácil	Fácil	Nem fácil nem difícil	Difícil	Muito difícil			MÍNIMO	MÉDIO	MÁXIMO	
1	2	5,0%	13	32,5%	23	57,5%	2	5,0%	0	0,0%	99
2	1	2,5%	10	25,0%	13	32,5%	14	35,0%	2	5,0%	74
3	2	5,0%	13	32,5%	17	42,5%	3	7,5%	5	12,5%	84
4	3	7,5%	13	32,5%	18	45,0%	6	15,0%	0	0,0%	93
5	0	0,0%	11	27,5%	15	37,5%	14	35,0%	0	0,0%	77
6	0	0,0%	7	17,5%	23	57,5%	10	25,0%	0	0,0%	77
7	4	10,0%	18	45,0%	13	32,5%	3	7,5%	2	5,0%	99
8	1	2,5%	14	35,0%	13	32,5%	11	27,5%	1	2,5%	83
9	1	2,5%	12	30,0%	18	45,0%	7	17,5%	2	5,0%	83
10	1	2,5%	5	12,5%	12	30,0%	13	32,5%	5	12,5%	96
11	1	2,5%	9	22,5%	16	40,0%	10	25,0%	2	5,0%	73
12	9	22,5%	15	37,5%	10	25,0%	6	15,0%	0	0,0%	107
13	5	12,5%	10	25,0%	18	45,0%	6	15,0%	1	2,5%	92
14	12	30,0%	9	22,5%	13	32,5%	5	12,5%	1	2,5%	106
15	4	10,0%	19	47,5%	12	30,0%	4	10,0%	1	2,5%	101
16	8	20,0%	19	47,5%	13	32,5%	0	0,0%	0	0,0%	115
17	1	2,5%	12	30,0%	15	37,5%	11	27,5%	1	2,5%	81
18	0	0,0%	5	12,5%	13	32,5%	16	40,0%	6	15,0%	57
19	2	5,0%	13	32,5%	16	40,0%	8	20,0%	1	2,5%	87
20	4	10,0%	19	47,5%	10	25,0%	5	12,5%	2	5,0%	98
							ESCORES ESPERADOS			ESCORE BRUTO OBTIDO (Seção 1)	ESCORE PONDERADO OBTIDO (Seção 1)
							MÍNIMO	MÉDIO	MÁXIMO	1738	34,72
							0	50	100		

O Produto - Documentação Online e Offline											
ASPECTO	ESCALA					NR	ESCORE OBTIDO	ESCORE ESPERADOS	ESCORE BRUTO OBTIDO (Seção 2)	ESCORE PONDERADO OBTIDO (Seção 2)	
	1	2	3	4	5						
	Muito fácil	Fácil	Nem fácil nem difícil	Difícil	Muito difícil			MÍNIMO	MÉDIO	MÁXIMO	
21	11	27,5%	17	42,5%	5	12,5%	7	17,5%	0	0,0%	112
22	6	15,0%	16	40,0%	11	27,5%	7	17,5%	0	0,0%	101
23	4	10,0%	11	27,5%	13	32,5%	6	15,0%	3	7,5%	81
24	3	7,5%	7	17,5%	25	62,5%	5	12,5%	0	0,0%	88
25	5	12,5%	7	17,5%	19	47,5%	9	22,5%	0	0,0%	88
26	3	7,5%	10	25,0%	7	17,5%	11	27,5%	6	15,0%	67
27	3	7,5%	12	30,0%	14	35,0%	5	12,5%	1	2,5%	81
28	1	2,5%	11	27,5%	15	37,5%	8	20,0%	0	0,0%	73
29	3	7,5%	13	32,5%	14	35,0%	3	7,5%	0	0,0%	88
30	0	0,0%	6	15,0%	21	52,5%	6	15,0%	3	7,5%	66
							ESCORES ESPERADOS			ESCORE BRUTO OBTIDO (Seção 2)	ESCORE PONDERADO OBTIDO (Seção 2)
							MÍNIMO	MÉDIO	MÁXIMO	847	56,47
							0	50	100		

Você e o Produto											
ASPECTO	ESCALA					NR	ESCORE BRUTO	ESCORE ESPERADOS	ESCORE BRUTO OBTIDO (Seção 3)	ESCORE PONDERADO OBTIDO (Seção 3)	
	1	2	3	4	5						
	Concordo totalmente	Concordo parcialmente	Nem concordo nem discordo	Discordo parcialmente	Discordo totalmente			MÍNIMO	MÉDIO	MÁXIMO	
31	11	27,5%	19	47,5%	2	5,0%	3	7,5%	5	12,5%	108
32	6	15,0%	18	45,0%	1	2,5%	9	22,5%	6	15,0%	89
33	3	7,5%	13	32,5%	10	25,0%	11	27,5%	3	7,5%	82
34	7	17,5%	9	22,5%	2	5,0%	6	15,0%	16	40,0%	93
35	5	12,5%	11	27,5%	11	27,5%	11	27,5%	2	5,0%	86
36	5	12,5%	11	27,5%	7	17,5%	12	30,0%	5	12,5%	79
37	14	35,0%	13	32,5%	2	5,0%	8	20,0%	3	7,5%	107
38	16	40,0%	11	27,5%	4	10,0%	6	15,0%	3	7,5%	92
39	6	15,0%	21	52,5%	3	7,5%	8	20,0%	2	5,0%	101
40	2	5,0%	16	40,0%	12	30,0%	6	15,0%	4	10,0%	86
41	15	37,5%	10	25,0%	6	15,0%	4	10,0%	5	12,5%	94
42	21	52,5%	6	15,0%	7	17,5%	6	15,0%	0	0,0%	98
43	10	25,0%	10	25,0%	9	22,5%	9	22,5%	1	2,5%	99
44	18	45,0%	9	22,5%	4	10,0%	7	17,5%	2	5,0%	96
45	10	25,0%	8	20,0%	7	17,5%	8	20,0%	5	12,5%	86
46	3	7,5%	12	30,0%	2	5,0%	9	22,5%	14	35,0%	99
47	6	15,0%	17	42,5%	6	15,0%	7	17,5%	4	10,0%	86
48	4	10,0%	11	27,5%	11	27,5%	11	27,5%	3	7,5%	78
49	15	37,5%	17	42,5%	2	5,0%	5	12,5%	1	2,5%	120
50	21	52,5%	10	25,0%	4	10,0%	5	12,5%	0	0,0%	127
							ESCORES ESPERADOS			ESCORE BRUTO OBTIDO (Seção 3)	ESCORE PONDERADO OBTIDO (Seção 3)
							MÍNIMO	MÉDIO	MÁXIMO	1635	51,29
							0	50	100		



ESCORES ESPERADOS					ESCORE BRUTO GLOBAL (Seções 1, 2 e 3)	ESCORE PONDERADO GLOBAL (Seções 1, 2 e 3)
MÍNIMO	MÉDIO	MÁXIMO				
0	50	100		4220	53,66	

Conforme descrito no capítulo anterior, os 24 itens do *DePerUSI* destinam-se à sondagem de 19 características (i) *físicas*, (ii) *relativas ao conhecimento e à experiência*; e (iii) *relativas à tarefa e ao trabalho*, possibilitando o “mapeamento” de universos amostrais de usuários envolvidos em diferentes contextos avaliatórios e facilitando sua caracterização. O Quadro 47 mostra que os itens 2, 3, 4 e 5 são destinados à sondagem de características *físicas*, enquanto os itens 1, 9, 10 a 14 e 18 a 24 permitem a aquisição de dados *relativos ao conhecimento e à experiência* do respondente. Quanto aos demais itens (6 a 8 e 15 a 17), estes possibilitam a aquisição de informações *relativas à tarefa e ao trabalho* do respondente.

As Figs. 59 a 72, apresentadas a seguir, detalham graficamente os resultados sumarizados no Quadro 47.

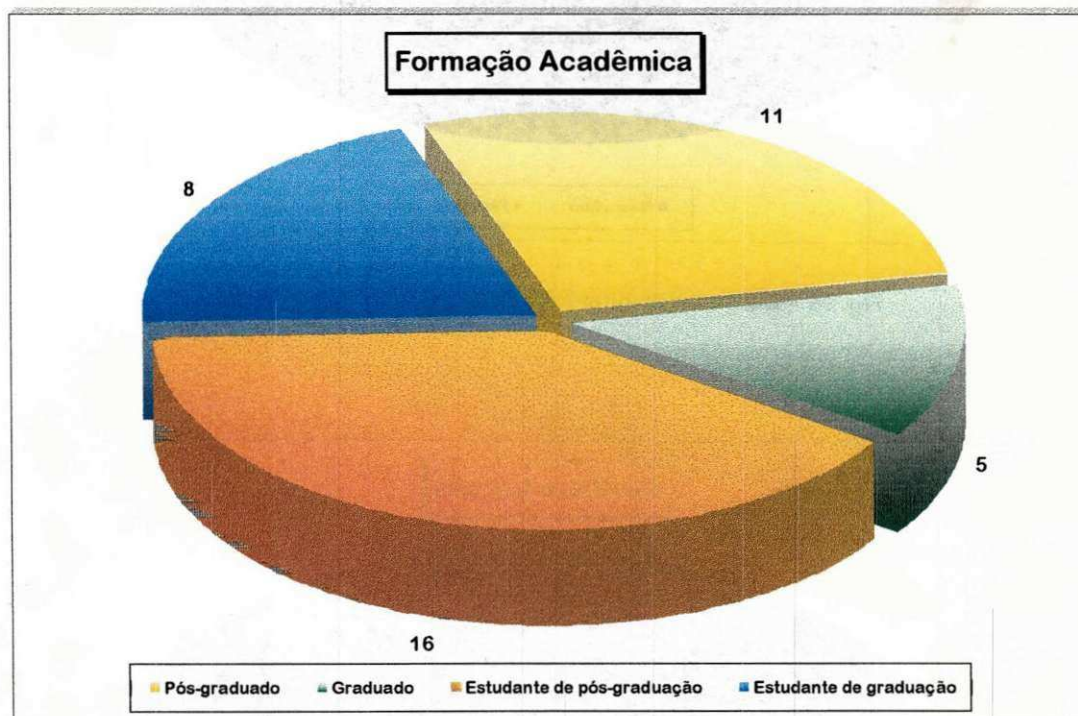


Fig. 57 – Distribuição numérica da *Formação Acadêmica* dos respondentes.

De acordo com a Fig. 57, a maior parte dos usuários do *MATLAB* que participaram do ensaio de usabilidade compunha-se de *estudantes de pós-graduação* (mestrado e doutorado), i.e. **16** indivíduos (**40%** do universo amostral), seguida de *profissionais pós-graduados* (mestres ou doutores) - atuantes como professores ou funcionários no Campus II da *UFPB* (**11** indivíduos ou **27,5%** do universo amostral), *estudantes de graduação* (**8** indivíduos ou **20%** do universo amostral) e de *indivíduos graduados*, atuantes ou não no Campus II da *UFPB* (**5** indivíduos ou **12,5%** do universo amostral). Dentre os respondentes, em sua maior parte do sexo masculino (Fig. 58), predominou o uso da mão direita (Fig. 59) e de corretivos visuais (Fig. 60).

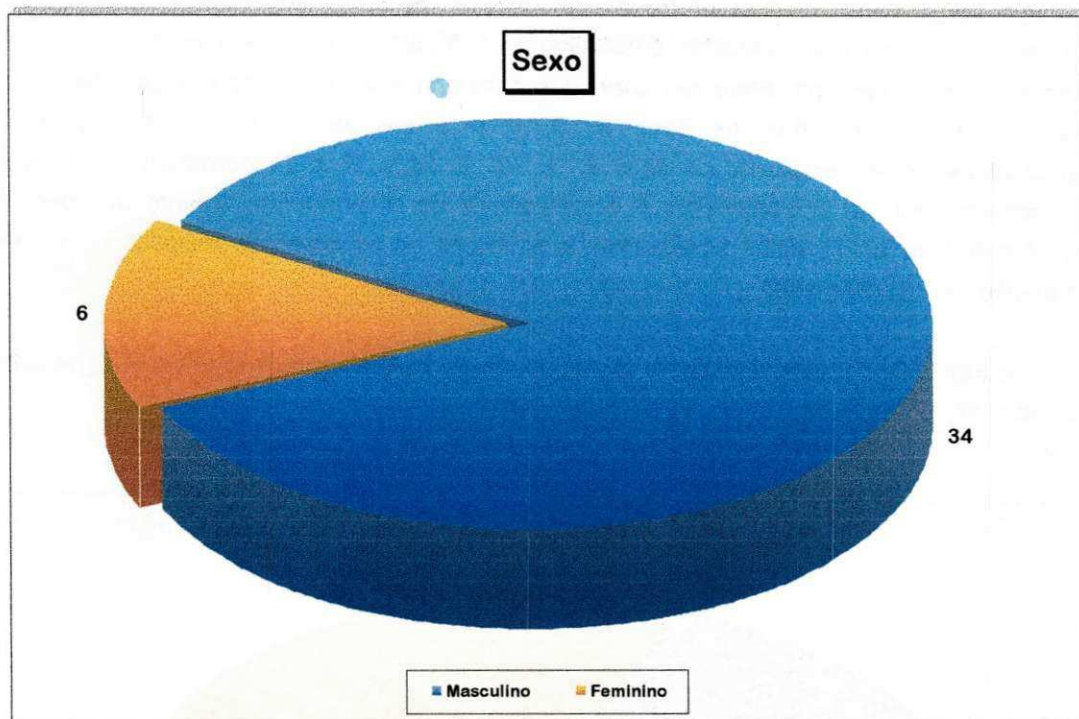


Fig. 58 – Distribuição numérica do Sexo dos respondentes.

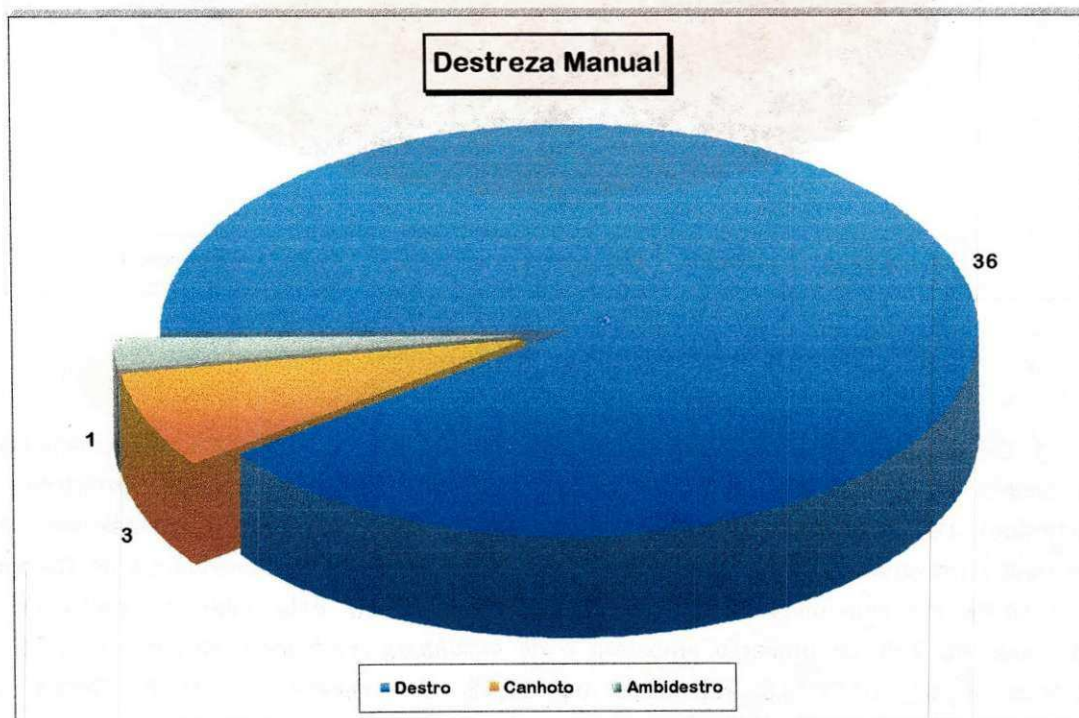


Fig. 59 – Distribuição numérica da *Destreza Manual* dos respondentes.

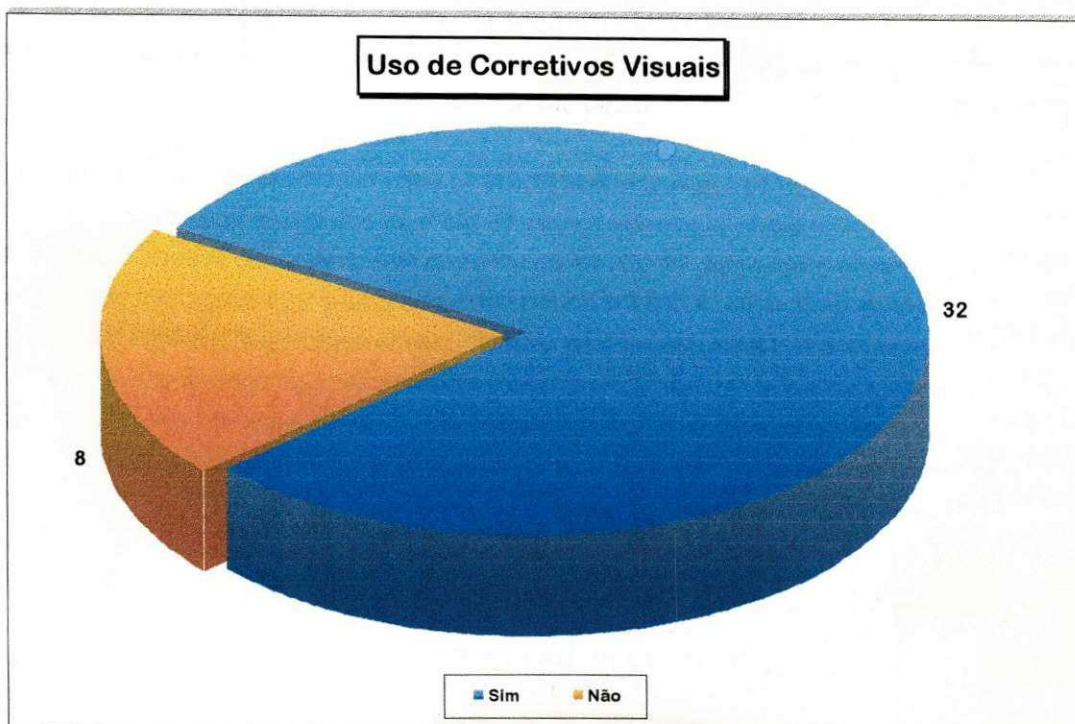


Fig. 60 – Distribuição numérica do *Uso de Corretivos Visuais* dentre os respondentes.

Quanto às faixas etárias, a Fig. 61 mostra que o maior número de respondentes (80% do universo amostral) pertencem, em ordem de representatividade, às faixas de 18 a 24 anos (18 indivíduos ou 45% do universo amostral) e de 25 a 34 anos (14 indivíduos ou 35% do universo amostral), seguidas das faixas de 35 a 45 anos (6 indivíduos ou 15% do universo amostral) e acima de 45 anos (2 indivíduos ou 5% do universo amostral).

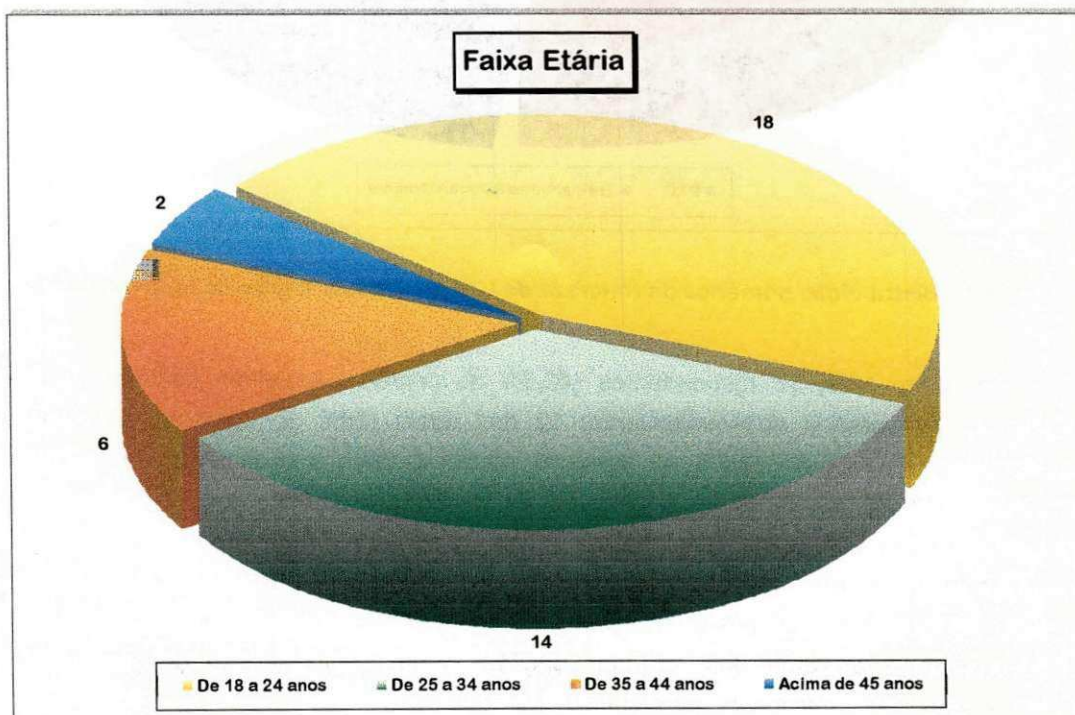


Fig. 61 – Distribuição numérica da *Faixa Etária* dos respondentes.

Conforme se observa no Quadro 47 e nas Figs. 62 e 63, nenhum respondente faz uso imediato dos utilitários do produto, quer para a prestação de serviços, execução de projetos, desenvolvimento de produtos ou treinamento em nível de extensão.

A Fig. 62 mostra que **23** respondentes (**57,5%** do universo amostral) afirmaram utilizar o produto para fins de *treinamento e aprendizagem* - **16** (**40%** do universo amostral) como suporte laboratorial em nível de graduação, **16** (**37,5%** do universo amostral) como suporte a disciplinas teóricas em cursos de graduação, **9** (**22,5%** do universo amostral) como suporte laboratorial em nível de pós-graduação e **5** (**12,5%** do universo amostral) como suporte a disciplinas teóricas em cursos de pós-graduação (vide Fig. 60). É importante ressaltar que, quer com suporte ao ensino ou ao aprendizado em nível de graduação ou pós-graduação, há professores e/ou estudantes que utilizam o produto simultaneamente em disciplinas teóricas e práticas, daí o montante dos percentuais exceder 100% do universo amostral.

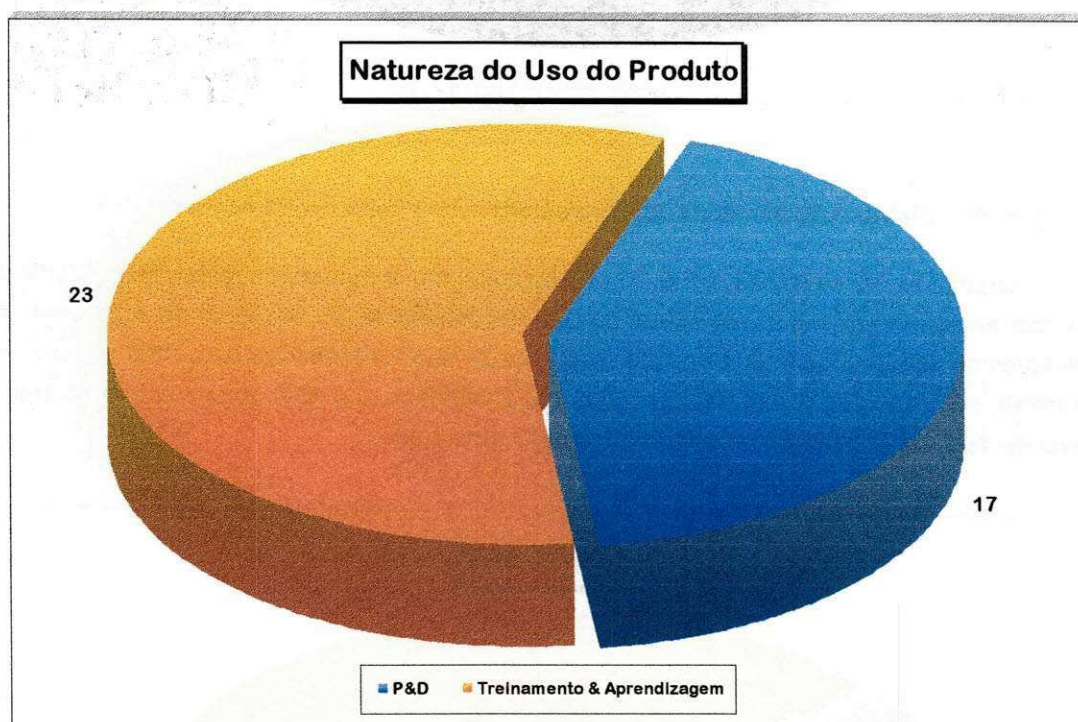


Fig. 62 – Distribuição numérica da *Natureza de Uso do Produto* dentre os respondentes.

Por outro lado, **17** respondentes (**42,5%** do universo amostral) declararam usar o produto em *pesquisa e desenvolvimento*, **12** dos quais (**30%** do universo amostral) em dissertações ou teses, **9** (**22,5%** do universo amostral) para a construção de modelos, **7** (**17,5%** do universo amostral) na verificação de modelos existentes e **4** (**10%** do universo amostral) em trabalhos de iniciação científica, conforme discriminado na Fig. 63. É importante ressaltar que, quer com suporte ao desenvolvimento de trabalhos de pesquisa em nível de graduação ou pós-graduação, também houve professores e/ou estudantes que assinalaram mais de uma das referidas opções, o que indica o uso do produto como suporte a diferentes iniciativas paralelas de pesquisa.

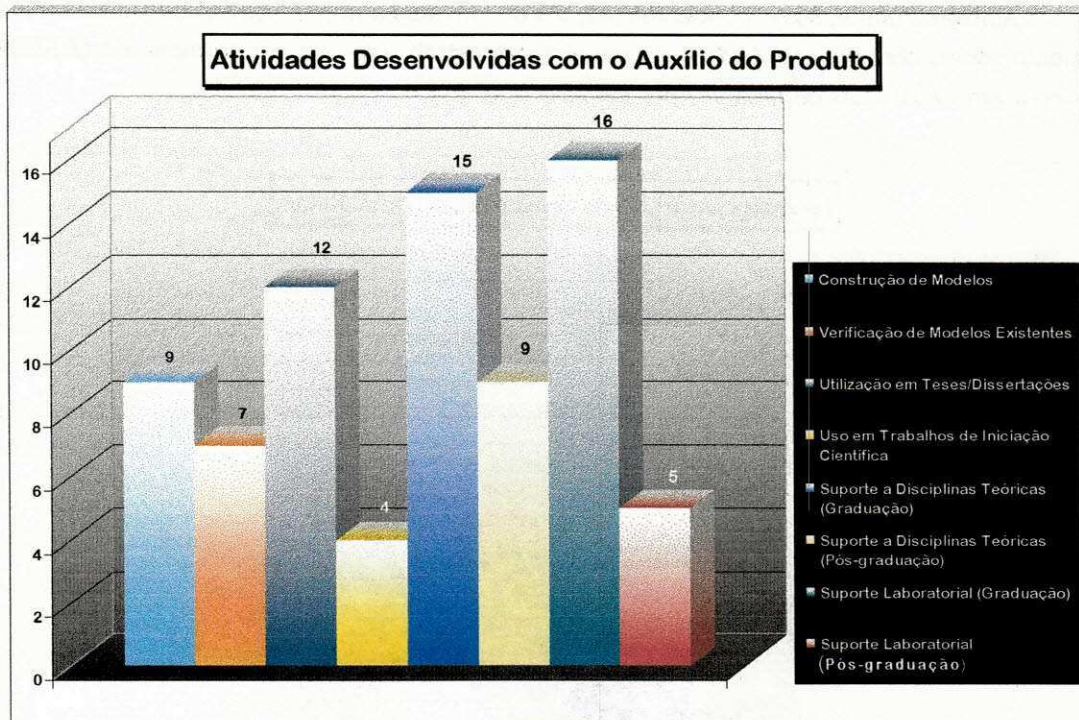


Fig. 63 – Distribuição numérica das *Atividades Desenvolvidas com o Auxílio do Produto* pelos respondentes.

No tocante aos itens 8, 9 e 12 do *DePerUSI*, todos os respondentes declararam utilizar versões do produto em computadores pessoais e ter experiência prévia tanto com sistemas computacionais quanto com o produto. A maioria dos respondentes (39 indivíduos ou 97,5% do universo amostral) afirmou vir usando sistemas computacionais há *mais de 1 ano*, enquanto o restante assinalou a opção *entre 3 meses e 1 ano* (Fig. 64).



Fig. 64 – Distribuição numérica do *Tempo de Uso de Sistemas Computacionais*.

Adicionalmente, **33** respondentes (**82,5%** do universo amostral) declararam usar sistemas computacionais *diariamente*, **4** (**10%** do universo amostral) *uma vez por semana* e **3** (**7,5%** do universo amostral) *mais de uma vez por semana* (Fig. 65).

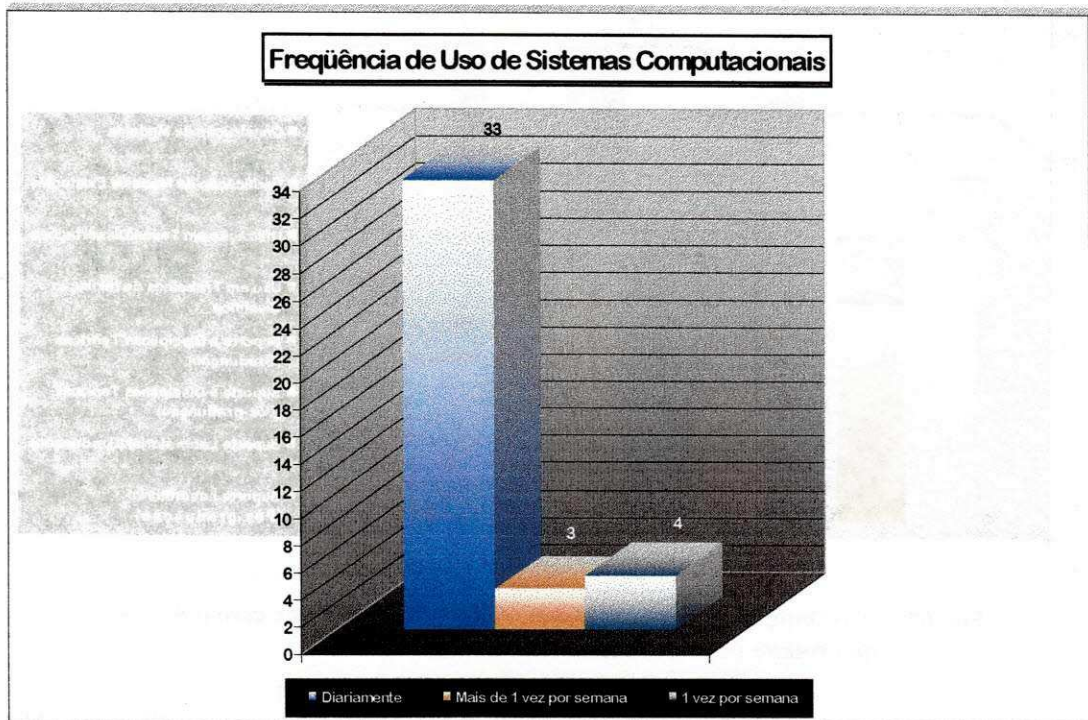


Fig. 65 – Distribuição numérica da *Frequência de Uso de Sistemas Computacionais*.

Por outro lado, a maior parte dos respondentes declarou usar o produto há *mais de 1 ano* (**22** indivíduos ou **55%** do universo amostral), segundo o Quadro 47. A este contingente, acrescentaram-se, em ordem decrescente, declarações de uso há *menos de 3 meses* (**11** indivíduos ou **27,5%** do universo amostral) e *entre 3 meses e 1 ano* (**7** indivíduos ou **17,5%** do universo amostral).

Intimamente relacionadas com estas cifras, as respostas ao item 14 do *DePerUSI* mostraram que, no tocante à *frequência de uso do produto*, os dois contingentes mais representativos do universo amostral têm usado o produto *mais de 1 vez por semana* (**12** indivíduos ou **30%** do universo amostral) e *1 vez por mês* (**9** indivíduos ou **22,5%** do universo amostral), respectivamente. Em terceiro lugar, se encontram os respondentes que têm usado o produto *diariamente*.

Observe-se que há as respostas aos itens 13 e 14 do *DePerUSI* permitem compreender melhor os resultados do *procedimento Tukey-Kramer*, sumariados graficamente através do Quadro 44, ao final da seção 7.4.2, nos quais predominam diferenças estatisticamente perceptíveis entre as categorias *extremas* de usuários de teste, i.e. *principiantes* e *experientes*. Obviamente, tais diferenças já eram esperadas, tendo em vista ocorrerem em diversos experimentos estatísticos similares ao ensaio de usabilidade conduzido nesta pesquisa. Entretanto, uma análise conjunta dos itens 13 e 14 mostra que embora a maioria dos respondentes tenha declarado usar o produto há mais de um ano, a maior parte não o tem usado diariamente. Eis porque as categorias de usuários mais numerosas do ensaio de usabilidade foram aquelas constituídas de usuários

principiantes e *intermediários*. Eis porque esta última categoria se mostrou tão heterogênea com relação às demais, incluindo indivíduos com perfis similares àqueles exibidos tanto pela categoria *principiantes* quanto pela categoria *experientes*. Graças a este fato, não foi possível identificar diferenças estatisticamente perceptíveis entre usuários *principiantes* e *intermediários* e entre estes e os usuários classificados como *experientes*.

As Figs. 66 e 67 ilustram, respectivamente, as distribuições numéricas referentes às respostas aos itens 13 e 14.

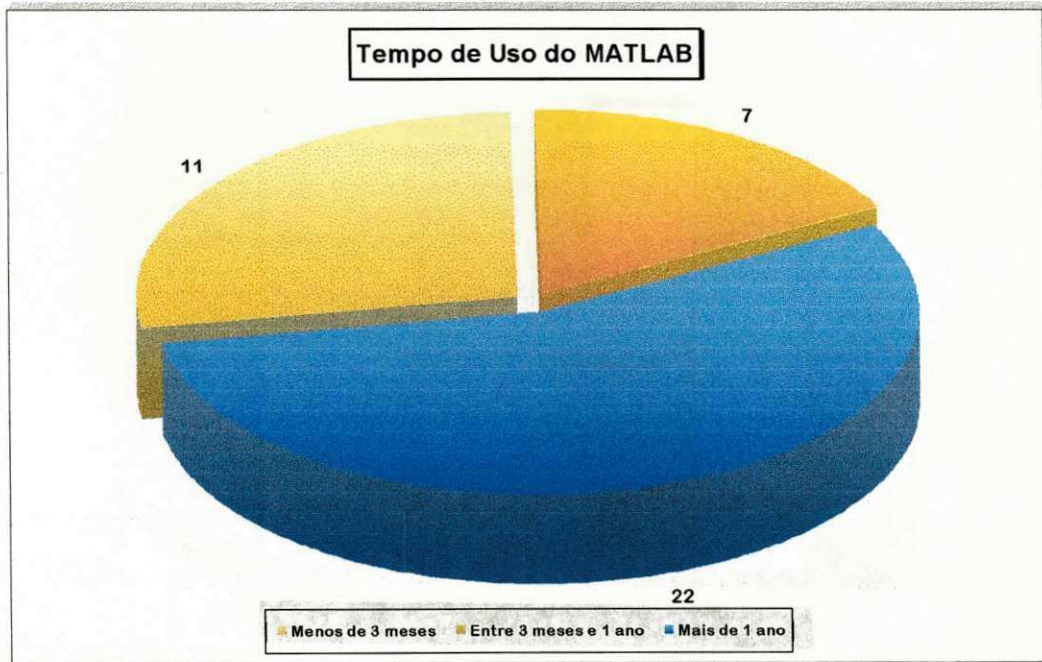


Fig. 66 – Distribuição numérica do *Tempo de Uso do Produto*.

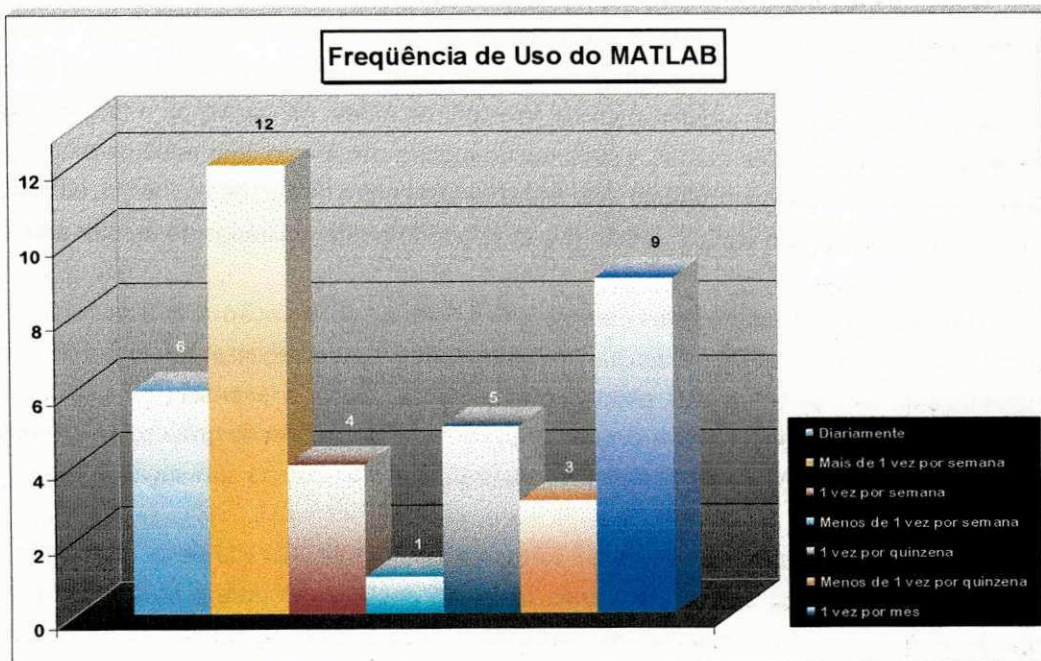


Fig. 67 – Distribuição numérica da *Frequência de Uso do Produto*.

Outro aspecto que influenciou na classificação dos usuários de teste foi a versão utilizada do produto, aspecto sondado através do item 15 do *DePerUSI*. A Fig. 68 ilustra as distribuições numéricas das respostas a este item, onde se observa que o maior número de respondentes (21 indivíduos ou 52,5% do universo amostral) declarou usar a versão 5.2, cuja interface, apesar de similar à versão avaliada em diversos aspectos, apresenta algumas diferenças significativas nos mecanismos de interação do usuário com os recursos oferecidos pelo produto.

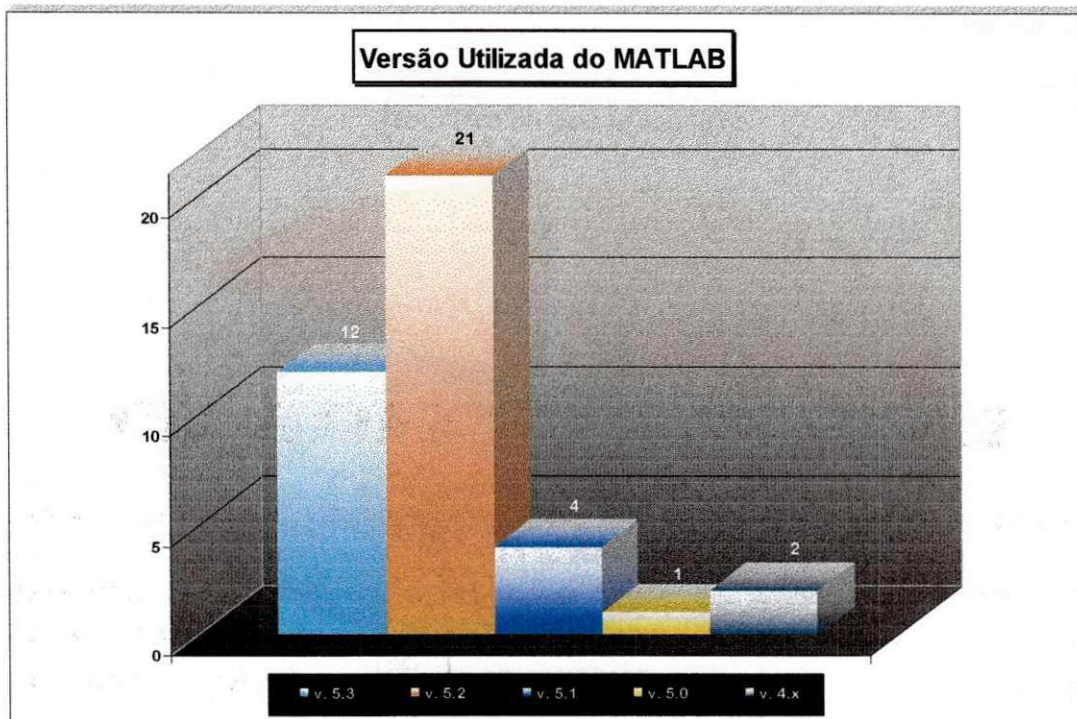


Fig. 68 – Distribuição numérica do aspecto *Versão Utilizada do MATLAB*.

Conforme anteriormente mencionado nas sub-seções referentes às falhas identificadas a partir da inspeção de conformidade e da mensuração do desempenho do usuário, o produto avaliado não privilegia inteiramente a experiência prévia do usuário com mecanismos de interação de outros produtos (e.g., os mecanismos de consulta da janela da ajuda) ou o faz, em alguns casos (e.g., similaridade dos nomes e sintaxes de alguns comandos com aqueles empregados pelo *MS-DOS*), de forma inconsistente. Isto se reproduziu entre as versões 5.2 e 5.3, dificultando em maior ou menor grau o uso da versão avaliada por diferentes usuários da versão anterior.

Um exemplo das dificuldades constatadas a partir da observação direta, da verbalização informal de ações pelos usuários e das análises retrospectivas dos registros em vídeo diz respeito à manipulação dos atributos dos objetos gráficos (cor, estilo e espessura de linhas e de marcadores) e textuais (tipo, estilo e cor de fontes) no interior da área do gráfico, assim como ao uso dos recursos de alteração da perspectiva do gráfico, discutidos na sub-seção 7.4.3.

Ao item 16 do *DePerUSI*, relativo à forma de ajuda mais utilizada pelos respondentes, a maioria (22 indivíduos ou 55% do universo amostral) declarou empregar a *ajuda de linha de comando*. A este número, seguiram-se, em ordem decrescente, as *consultas orais a outros usuários* (7 indivíduos ou 17,5% do universo amostral), as *consultas a manuais* (6 indivíduos ou 15% do universo amostral), as *consultas à HELPWIN* (4 indivíduos ou 10% do universo amostral) e

à HELPDESK (Internet) (1 indivíduo ou 2,5% do universo amostral), conforme ilustra a Fig. 69.

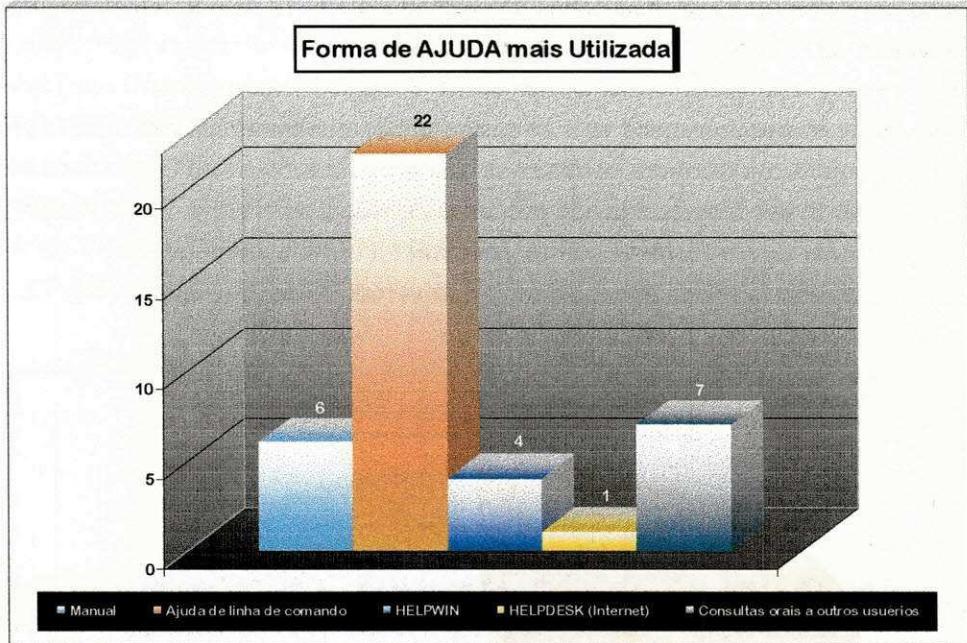


Fig. 69 – Distribuição numérica da *Forma de AJUDA mais Utilizada*, segundo os respondentes.

Embora as tarefas de teste não tenham contemplado o uso de caixas de ferramentas (*toolboxes*) específicas, o item 17 do *DePerUSI* sondou seu uso em atividades cotidianas dos usuários de teste. Das 24 caixas de ferramentas listadas¹⁰⁰, apenas 11 foram apontadas pelos respondentes, de acordo com o Quadro 47 ou com a Fig. 70.

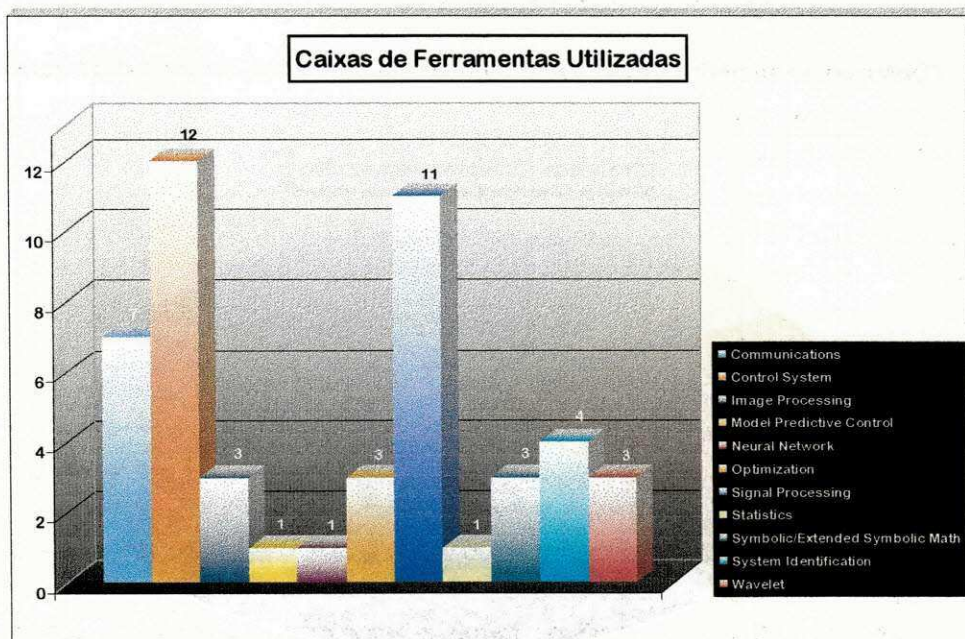


Fig. 70 – Distribuição numérica das *Caixas de Ferramentas (Toolboxes)*.

¹⁰⁰ Foram excluídas da lista, após a etapa de recrutamento dos participantes do ensaio de usabilidade, as *toolboxes Database e Financial*, em função de suas áreas de concentração.

Sobressaíram nesta sondagem, em ordem decrescente, as *toolboxes Control System* (12 indivíduos ou 30% do universo amostral), *Signal Processing* (11 indivíduos ou 27,5% do universo amostral) e *Communications* (7 indivíduos ou 17,5% do universo amostral), conforme ilustra a Fig. 70.

No que diz respeito aos itens 18 a 24 do *DePerUSI*, que tratam da experiência prévia do respondente com produtos similares, verificou-se que, embora a maioria (21 indivíduos ou 52,5% do universo amostral) tenha declarado não possuir experiência prévia (Fig. 71), a totalização das respostas evidenciou que o restante dos respondentes (19 indivíduos ou 47,5% do universo amostral) tem experiência prévia com pelo menos um produto, discriminado na Fig. 72.

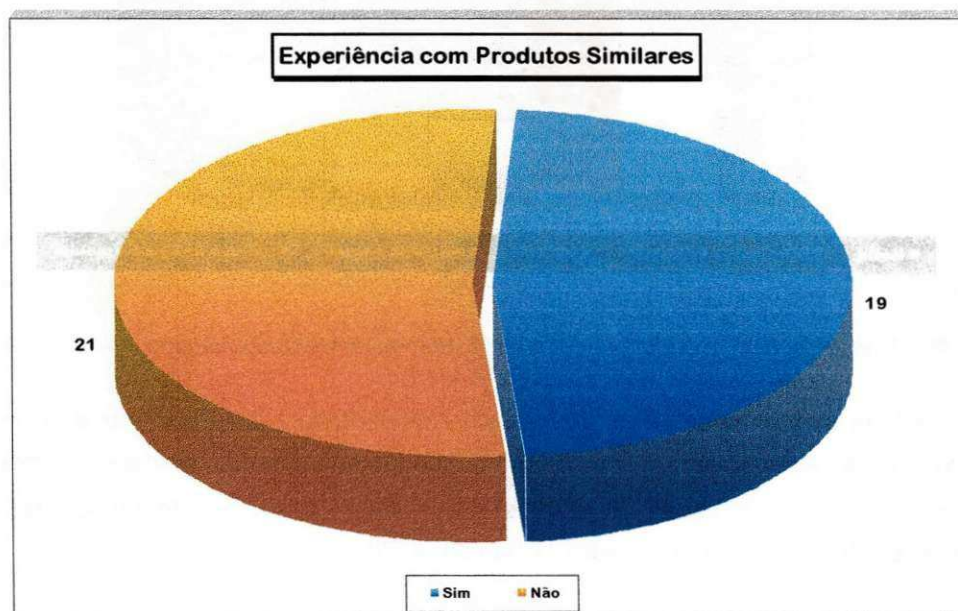


Fig. 71 – Distribuição numérica da *Experiência Prévia com Produtos Similares* dos respondentes.

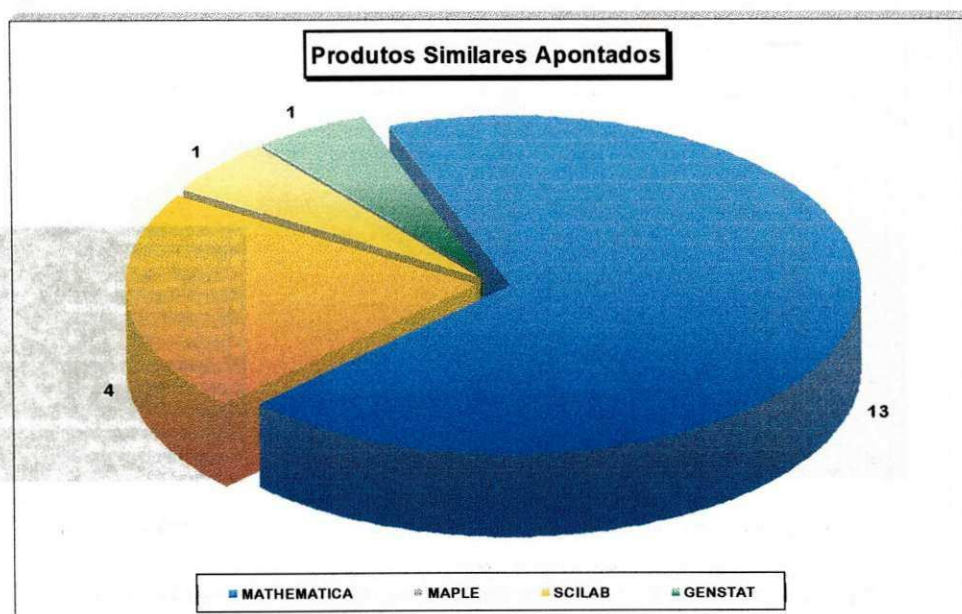


Fig. 72 – Distribuição numérica dos produtos similares, segundo os respondentes.

Dos 19 respondentes (47,5% do universo amostral) que declararam ter experiência prévia com produtos similares ao *MATLAB*, a maioria (13 indivíduos ou 32,5% do universo amostral) indicou o *MATHEMATICA*, 4 (10% do universo amostral) indicaram o *MAPLE*, 1 (2,5% do universo amostral) o *SCILAB* e 1 (2,5% do universo amostral) o *PTOLEMY* (Fig. 72).

Finalmente, dentre aqueles que declararam ter experiência prévia com um produto similar ao *MATLAB*, 8 (20% do universo amostral) assinalaram *menos de 3 meses*, 7 (17,5% do universo amostral) *entre 3 meses e 1 ano* e 4 (10% do universo amostral) *mais de 1 ano* de uso (Fig. 73).

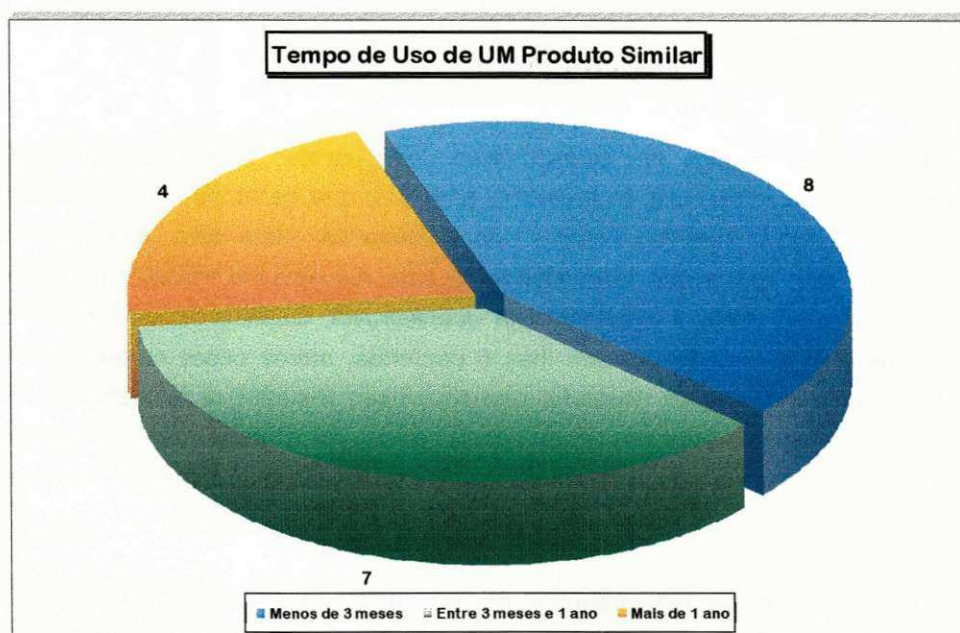


Fig. 73 – Distribuição numérica do *Tempo de Uso de UM Produto Similar* ao *MATLAB*.

Diante do exposto ao longo desta sub-seção, pode-se afirmar, em suma, que o respondente típico do *DePerUSI* neste ensaio de usabilidade, usuário de uma das versões do produto avaliado, apresenta as seguintes características:

- ✦ É *estudante de pós-graduação* (mestrado ou doutorado) ou *graduado* (mestre ou doutor), predominantemente do sexo *masculino*, idade *entre 18 e 35 anos*, *destro* e *portador de corretivos visuais* (óculos ou lentes de contato);
- ✦ Possui experiência computacional prévia de *mais de 1 ano* e utiliza *diariamente* computadores pessoais (PC), predominantemente em atividades de *pesquisa e desenvolvimento*;
- ✦ Utiliza o produto como suporte em teses/dissertações ou em atividades relativas a disciplinas teóricas, quer como instrutor ou como aprendiz;
- ✦ Possui experiência prévia de *mais de 1 ano* com o produto avaliado, utilizando *mais de 1 vez por semana* a versão v. 5.2 ou a versão avaliada do produto (v. 5.3);

- ✦ Utiliza preferencialmente a *ajuda de linha de comando*, provavelmente devido à influência dos instrutores com experiência de uso das versões mais antigas do produto (para *MS-DOS*), nas quais esta forma de ajuda era a única existente;
- ✦ Utiliza recursos da caixa de ferramentas *Control System*, *Signal Processing* ou *Communications*, o que sugere se tratar de um professor, funcionário ou estudante do curso de Engenharia Elétrica; e
- ✦ Possui experiência com um produto similar, o *Mathematica*, de *menos de 3 meses*, na qualidade de instrutor ou aprendiz, via de regra devido ao referido produto ser utilizado como suporte a disciplinas teóricas e laboratoriais.

7.4.2 Resultados da Análise do OpUS

A computação de respostas em questionários com escalas numéricas “ancoradas” por termos de semântica diferencial é dicótoma, uma vez que se atribui *1 ponto* a cada item assinalado da escala e *0 pontos* àqueles não assinalados. Além disto, pode-se atribuir *+1 ponto* quando o item assinalado corresponder a uma resposta *favorável*, em contraposição a atribuição *-1 ponto*, caso a resposta seja *desfavorável*. Quando os itens são ordenados na escala conforme a importância que lhes é conferida, outros pesos numéricos diferentes de *0* e *1* podem ser utilizados. Todavia, segundo Aiken [Aike96], a ponderação diferencial dos itens exerce usualmente pouca influência sobre a confiabilidade e a validade de instrumentos de sondagem com um grande número de itens, e.g. 50 itens, como é o caso do *OpUS*.

Exceto quando os respondentes de um questionário tiverem sido instruídos a assinalar um número especificado de itens, diferentes indivíduos poderão assinalar diferentes números de itens, quer por não se sentirem seguros a emitir uma opinião sobre o aspecto sondado, por não se acharem aptos a opinar ou, simplesmente, por se recusarem a fazê-lo, como ocorreu com os itens 27 a 29, que não foram preenchidos por $\frac{1}{4}$ dos respondentes da categoria *princiantes* (4 indivíduos, 25% dos indivíduos da classe ou 10% do universo amostral). Tal variação no número de itens assinalados poderá exercer um efeito significativo sobre a determinação dos escores, de modo a tornar necessário o uso de alguma forma de agrupamento ou de controle estatístico.

O Quadro 48 resume a totalização das respostas dos usuários de teste aos itens do instrumento de sondagem da opinião do usuário sobre o *MATLAB v. 5.3.0*, tendo sido construído de modo a manter a estrutura do *OpUS*, com o propósito de facilitar a associação dos 50 itens às opções assinaladas nas duas escalas semântico-numéricas de 5 pontos – *muito fácil a muito difícil* e *concordo totalmente a discordo totalmente*. Os resultados mais expressivos foram contornados com círculos coloridos, obedecendo à seguinte convenção: (i) *azul*, associado às respostas positivas mais expressivas; (ii) *cinza escuro*, associado às respostas neutras mais expressivas; e (iii) *vermelho*, associado às respostas negativas mais expressivas.

Foram adicionadas à estrutura do *OpUS* três colunas, a saber: (i) *NR*, que contém, em valores absolutos e percentuais, informações sobre o número de respondentes que se omitiram de opinar sobre alguns dos itens do instrumento de sondagem; (ii) *ESCORE BRUTO*, contendo os resultados não ponderados da computação de cada item; e (iii) uma representação gráfica em cores, associadas às faixas de escores, às quais foram atribuídos os referenciais semânticos *BAIXO*, *RAZOÁVEL*, *BOM* e *MUITO BOM*, conforme a legenda apresentada no canto inferior esquerdo do Quadro 48. Ao final de cada seção do *OpUS*, são apresentados os escores esperados (mínimo, máximo e médio) e os escores bruto e ponderado obtidos por seção.

Todavia, é necessário que se observe que, embora as opções positivas tenham sido posicionadas à esquerda e as negativas à direita de cada uma das escalas consideradas, as afirmativas associadas a cada item podem tornar as respostas à esquerda negativas e vice-versa, como ocorre nos itens 34, 38 e 41 a 48, em virtude das declarações formuladas terem sido expressas negativamente.

A computação de várias categorias, segmentos ou “âncoras” numéricas, semânticas ou semântico-numéricas ocorre mais facilmente se lhes forem atribuídos números inteiros sucessivos. Nos casos de escalas *unipolares*, aconselha-se atribuir um escore de 0 à categoria mais baixa da escala (e.g., 0 para a posição *muito difícil* da escala semântica *muito fácil-muito difícil* do *OpUS*) e um escore de $c-1$ (onde c é o número de categorias) à categoria mais elevada (e.g., 4 para a posição *muito fácil* da escala semântica *muito fácil-muito difícil* do *OpUS*). Valores próximos de 0 indicarão, portanto, baixos índices do aspecto sondado, ao passo que valores próximos de $c-1$ altos índices do referido aspecto.

Quando vários itens ou escalas em um questionário mensuram a mesma variável, deve-se determinar um escore composto, a partir da adição dos valores computados em tais itens. Nos casos em que se utiliza escalas *bipolares*, sugere-se a adoção de um número ímpar de categorias (c), às quais se deve associar inteiros consecutivos variando de $(1-c)/2$, para a categoria mais baixa, até $(c-1)/2$, para a categoria mais elevada, onde c é o número de categorias da escala [Aike96].

Tome-se, como exemplo, a escala de 5 pontos adotada no *OpUS* (cujos índices numéricos variam de 1 a 5) que, conforme esta sugestão, corresponderá a variações de -2 a +2, tendo como centro o índice 3, associado às âncoras semânticas *nem fácil nem difícil* e *nem concordo nem discordo*, as quais coincidem com o centro da escala. A nova escala de escores, variando de -2 a +2, fornecerá informações sobre a magnitude e a direção do conjunto de valores computados.

Todavia, ao discutir o *método da somatória de valores* (*method of summated ratings*), concebido por Likert [Like32, in Aike96], Aiken [Aike96] afirma que, para uma escala típica de 5 pontos (e.g., a escala adotado no *OpUS*), as respostas aos itens do questionário expressos positivamente devem ser computadas como segue: 0 para a posição mais desfavorável da escala (e.g., *discordo totalmente*), 1 para a posição parcialmente desfavorável (e.g., *discordo parcialmente*), 2 para a posição central (e.g., *indeciso* ou *não concordo nem discordo*), 3 para a posição parcialmente favorável (e.g., *concordo*

parcialmente) e 4 para a posição mais favorável (e.g., *concordo totalmente*). As respostas aos itens 1 a 33, 35 a 37, 39 a 40 e 49 a 50 do *OpUS* foram computados de acordo com esta afirmativa, resultando nos escores da última coluna do Quadro 48, correspondentes aos itens acima referidos.

Por outro lado, para as respostas aos itens do questionário expressos negativamente, deve-se inverter a ordem dos pesos atribuídos aos itens no processo de computação das respostas, i.e., 0 para a posição mais favorável da escala, 1 para a posição parcialmente favorável, 2 para a posição central, 3 para a posição parcialmente desfavorável e 4 para a posição mais desfavorável da escala. Os itens 34, 38 e 41 a 48 foram computados deste modo, o que resultou nos escores da última coluna do Quadro 48 para aos itens acima referidos.

Portanto, o escore total do conjunto de itens do tipo Likert associado a cada instrumento de sondagem, computado segundo o *método da somatória de valores* (*method of summated ratings*), será dado pela expressão:

$$E_i = \sum_{j=1}^n P_{ij}, \quad (9)$$

onde P_{ij} representa o peso numérico atribuído pelo indivíduo i ao item j e n representa o número de itens do instrumento de sondagem.

A extensão da expressão (9) para todos os respondentes resulta na computação do escore global dos instrumentos de sondagem recebidos, i.e.

$$E_G = N \cdot E_i = N \sum_{j=1}^n P_{ij}, \quad (10)$$

onde N representa o número de instrumentos de sondagem recebidos, ou seja, o número de respondentes.

Vale a pena observar que o melhor resultado esperado de um instrumento de sondagem dessa natureza corresponderá às respostas mais favoráveis a todos os aspectos sondados, i.e., a atribuição do peso mais elevado da escala para todos os itens. Considerando o *OpUS* como exemplo, a atribuição do peso máximo 4 a todos os itens implicará a obtenção de um escore total bruto igual a $4 \times 50 = 200$ para cada questionário. Na computação do escore global relativo ao universo amostral considerado, o montante será, portanto, $40 \times 200 = 8000$. Observe-se os montantes brutos referentes a cada seção do *OpUS* e o montante global do universo amostral considerado, sumariados no Quadro 49.

Analogamente ao melhor resultado, o pior resultado esperado equivalerá às respostas menos favoráveis a todos os aspectos sondados, i.e., a atribuição do peso mais baixo da escala a todos os itens do questionário. Independentemente do número de posições da escala de Likert considerada, o escore bruto obtido neste caso será nulo.



ITEM	ESCORES DA SEÇÃO 3							
	PRINCIPIANTES		INTERMEDIÁRIOS		EXPERIENTES		TOTAL	
	BRUTO	PONDERADO	BRUTO	PONDERADO	BRUTO	PONDERADO	BRUTO	PONDERADO
01	35	2,19	38	2,38	22	2,75	95	2,38
02	25	1,56	30	1,88	19	2,38	74	1,85
03	31	1,94	36	2,25	17	2,13	84	2,10
04	36	2,25	40	2,50	17	2,13	93	2,33
05	34	2,13	26	1,63	17	2,13	77	2,14
06	28	1,75	33	2,20	16	2,00	77	2,03
07	41	2,56	39	2,44	19	2,38	99	2,48
08	31	1,94	34	2,13	18	2,25	83	2,08
09	34	2,13	30	1,88	19	2,38	83	2,08
10	25	1,92	19	1,19	12	1,71	56	1,40
11	29	1,81	30	1,88	14	2,00	73	1,83
12	44	2,75	44	2,75	19	2,38	107	2,68
13	35	2,19	40	2,50	17	2,13	92	2,30
14	44	2,75	42	2,63	20	2,50	106	2,65
15	38	2,38	44	2,75	19	2,38	101	2,53
16	44	2,75	46	2,88	25	3,13	115	2,88
17	33	2,06	31	1,94	17	2,13	81	2,03
18	20	1,25	23	1,44	14	1,75	57	1,43
19	37	2,31	32	2,00	18	2,25	87	2,18
20	43	2,69	36	2,25	19	2,38	98	2,45
	687	54,18	693	54,31	358	56,65	1738	54,72

ITEM	ESCORES DA SEÇÃO 3							
	PRINCIPIANTES		INTERMEDIÁRIOS		EXPERIENTES		TOTAL	
	BRUTO	PONDERADO	BRUTO	PONDERADO	BRUTO	PONDERADO	BRUTO	PONDERADO
21	44	2,75	49	3,06	19	2,38	112	2,80
22	38	2,38	43	2,69	20	2,50	101	2,53
23	28	2,15	37	2,85	16	2,00	81	2,19
24	32	2,00	36	2,25	20	2,50	88	2,20
25	28	1,75	40	2,50	20	2,50	88	2,20
26	19	1,46	33	2,20	15	1,88	67	1,81
27	27	2,25	35	2,33	19	2,38	81	2,31
28	23	1,92	33	2,06	19	2,38	75	2,14
29	27	2,25	40	2,50	21	2,63	88	2,51
30	25	1,92	28	1,75	13	1,63	66	1,83
	291	52,34	374	59,94	182	56,88	847	56,47

ITEM	ESCORES DA SEÇÃO 3							
	PRINCIPIANTES		INTERMEDIÁRIOS		EXPERIENTES		TOTAL	
	BRUTO	PONDERADO	BRUTO	PONDERADO	BRUTO	PONDERADO	BRUTO	PONDERADO
31	40	2,50	45	2,81	23	2,88	108	2,70
32	34	2,13	37	2,31	18	2,25	89	2,23
33	33	2,06	28	1,75	21	2,63	82	2,05
34	36	2,25	33	2,06	16	2,00	95	2,38
35	43	2,69	32	2,00	21	2,63	86	2,15
36	24	1,50	33	2,06	22	2,75	79	1,98
37	37	2,31	44	2,75	26	3,25	107	2,68
38	16	1,00	19	1,19	14	1,75	49	1,23
39	40	2,50	37	2,31	24	3,00	101	2,53
40	34	2,13	33	2,06	19	2,38	86	2,15
41	15	0,94	24	1,50	15	1,88	54	1,35
42	20	1,25	9	0,56	9	1,13	38	0,95
43	22	1,38	23	1,53	14	1,75	59	1,51
44	11	0,69	19	1,27	16	2,00	46	1,15
45	20	1,25	26	1,73	20	2,86	66	1,74
46	32	2,00	46	3,07	21	2,63	99	2,48
47	18	1,13	28	1,75	20	2,50	66	1,65
48	25	1,56	32	2,00	21	2,63	78	1,95
49	43	2,69	49	3,06	28	3,50	120	3,00
50	50	3,13	48	3,00	29	3,63	127	3,18
	593	46,33	645	50,71	397	62,42	1635	51,29

SOMATÓRIOS DOS ESCORES DAS 3 SEÇÕES							
PRINCIPIANTES		INTERMEDIÁRIOS		EXPERIENTES		TOTAL	
BRUTO	PONDERADO	BRUTO	PONDERADO	BRUTO	PONDERADO	BRUTO	PONDERADO
1571	50,61	1712	53,97	937	59,01	4220	53,66

A indecisão ou neutralidade completa do respondente será indicada pela atribuição do peso central da escala de Likert a todos os itens de um instrumento de sondagem. Para o *OpUS*, por exemplo, no qual o peso da posição central é igual a 2, o escore total bruto corresponderá a $2 \times 50 = 100$, para cada questionário computado. No que diz respeito ao escore global associado ao universo amostral considerado, o montante será, portanto, $40 \times 100 = 4000$.

A ponderação da faixa bruta abrangida pelos resultados extremos possíveis (i.e., o pior e o melhor resultados) para a faixa de 0 a 100 facilita a análise à primeira vista tanto do escore total dos n itens de um único instrumento de sondagem com escala de Likert, quanto do escore global resultante da computação das respostas aos n itens do referido instrumento pelos N respondentes do universo amostral considerado. Os resultados apresentados nas células das últimas linhas de cada seção do *OpUS* e na última linha do Quadro 48, assim como o sumário dos escores, feito no Quadro 49, respaldam a afirmação acima discutida.

7.4.2.1 Análise Preliminar dos Escores Obtidos

Como se pode observar no Quadro 49, a listagem dos escores por seção e por categoria de usuário do produto avaliado possibilita a visualização mais detalhada da opinião dos respondentes do *OpUS* sobre cada um dos temas investigados - (i) o uso das facilidades oferecidas pelo produto e a navegação do usuário por essas facilidades mediante a interface; (ii) a documentação *online* e *offline* fornecida pelo desenvolvedor do produto avaliado, no que concerne à localização, ao acesso e à compreensão das informações de interesse; e (iii) a relação do usuário com o produto, em um nível mais pessoal do que na sondagem dos dois primeiros temas.

A análise dos resultados obtidos para cada seção, por categoria de usuário, resulta na constatação de que os escores ponderados crescem da categoria *principiantes* para a categoria *experientes*, exceto na Seção 2 (**O Produto – Documentação *Online* e *Offline***), onde o escore associado à categoria *intermediários* é superior ao da categoria *experientes*. Esta ordenação se mantém com relação aos escores globais ponderados das categorias consideradas. Conforme se verificou com estes resultados, quanto maior o tempo e a frequência de uso do produto, mais favoráveis foram os respondentes na pontuação dos aspectos investigados do produto.

Esta constatação é consonante com os registros de outros autores (e.g. Dix *et al.* [Dix98], Constantine e Lockwood [Cons99]) sobre o comportamento de usuários experientes, no que diz respeito à incumbência de encontrar soluções alternativas para diversos problemas de usabilidade da interface. Tal postura implica usualmente expectativas mais modestas desta categoria de usuários quanto ao processo interativo com o produto e, por conseguinte, maior “condescendência” nas opiniões sobre diferentes aspectos do referido processo.

No tocante à Seção 1 (**O Produto – Uso e Navegação**), verifica-se que os escores relativos às categorias *principiantes* (54,18) e *intermediários* (54,31) são praticamente iguais e apenas ligeiramente inferiores ao escore relativo à categoria *experientes* (56,65). Contudo, esta

constatação merece um comentário. Durante a totalização das respostas dos 40 questionários em uma única planilha do *MS Excel 2000*, foram criadas planilhas auxiliares para a totalização das respostas de cada categoria, as quais mostraram que **50%** das respostas da categoria *principiantes* e **45%** da categoria *intermediários* aos itens da Seção 1 predominaram na posição central da escala de Likert de 5 pontos, contra apenas **25%** das respostas da categoria *experientes* predominantes na mesma posição da escala.

Quanto ao restante das respostas, verificou-se que tenderam moderadamente para o lado esquerdo da escala (opiniões favoráveis), com a maior tendência sendo registrada na planilha da categoria *experientes*. Os registros em vídeo e os comentários informais, principalmente de indivíduos desta última categoria, explicam a obtenção dos escores acima mencionados: foram registradas reações de indecisão e insegurança durante o preenchimento do *OpUS* por *principiantes* e *intermediários*, traduzidas por deslocamentos a esmo do *mouse* por sobre as opções da escala ("passeio aleatório"), questionamentos ao avaliador sobre tópicos irrelevantes às respostas, assinalações desordenadas de itens da seção, etc. Por outro lado, mais de **50%** dos respondentes da categoria *experientes* verbalizaram suas opções, comentando-as, o que enriqueceu sensivelmente o processo de sondagem.

A Seção 2 foi a que apresentou maior número de omissões nas respostas, tendo sido registradas pelo menos **3** abstinências em **60%** dos 10 itens, por parte dos respondentes *principiantes*, **1** abstinência em **40%** dos 10 itens, por parte dos *intermediários* e pelo menos **1** abstinência em **30%** dos itens, por parte dos *experientes*. Contrariamente as demais seções, a maior parte dos comentários verbais foi proveniente dos respondentes da categoria *principiantes*, traduzindo, em essência, a familiaridade insuficiente com os mecanismos de ajuda *online* e *offline* do produto, condição necessária para um posicionamento seguro sobre os aspectos sondados.

Por sua vez, a computação dos escores referentes à Seção 3 apresentou proporções similares àquelas da Seção 1, com os escores relativos às categorias *principiantes* (**46,33**) e *intermediários* (**50,71**) bem próximos entre si e inferiores ao escore relativo à categoria *experientes* (**62,42**). Mais uma vez, a análise das planilhas relativas a cada categoria de respondentes, aliada a análises dos registros em vídeo e dos comentários informais, conduziram a constatações similares àquelas comentadas para a Seção 1. Além do mais, estes resultados são consonantes com aqueles obtidos a partir da mensuração do desempenho do usuário durante o uso do produto avaliado, no que diz respeito às similaridades apresentadas pelas categorias *principiantes* e *intermediários*, não diferenciáveis estatisticamente.

Em suma, todas as categorias se mostraram, grosso modo, favoráveis ao produto, conforme se observa na última linha do Quadro 49. Entretanto, também se observa claramente que os escores totais referentes às categorias *intermediários* e *experientes* são moderadamente favoráveis ao produto (**53,97** e **59,01**, respectivamente), enquanto o escore total relativo à categoria *principiantes* reflete uma posição praticamente neutra (**50,61**) de seus integrantes quanto ao produto.

Por fim, este último resultado traduz a indecisão e a falta de familiaridade da categoria *principiantes* com os aspectos sondados, para os quais alguns dos integrantes da referida

categoria se mostraram omissos. A computação do número de omissões ao preenchimento de itens pelas diferentes categorias de respondentes, aliada aos comentários informais durante o preenchimento do *OpUS* e às análises dos registros em vídeo das sessões de teste, reforçam esta afirmativa.

7.4.2.2 Análise Estatística dos Indicadores Quantitativos

Similarmente à análise preliminar dos indicadores objetivos registrados nos testes de usabilidade que precederam a aplicação do *OpUS*, a análise do Quadro 49, apresentada na sub-seção anterior, teve o propósito evidenciar diferenças entre os escores registrados em função das categorias de respondentes do instrumento de sondagem utilizado nesta pesquisa.

Adicionalmente, fez-se necessário um processamento estatístico destes resultados, com o mesmo propósito da análise dos indicadores objetivos obtidos, a partir do teste *F ANOVA fator único*: verificar se as diferenças registradas na análise preliminar eram estatisticamente aceitáveis. Neste caso, a variável analisada foi o escore por seção e por categoria de respondente.

A hipótese nula de nenhuma diferença nas médias aritméticas dos escores relativos às diferentes categorias de respondentes ($H_0: \mu_{EP} = \mu_{EI} = \mu_{EE}$) foi testada em relação à alternativa de haver diferenças entre as referidas médias ($H_A: \text{nem todas as médias são iguais}$) onde μ_{EP} , μ_{EI} e μ_{EE} representam as médias aritméticas dos escores registrados, respectivamente, para as categorias *principiantes*, *intermediários* e *experientes*.

Vale a pena lembrar que a hipótese alternativa não representa precisamente a negação da hipótese nula, visto que, similarmente à mensuração do desempenho do usuário durante o uso do produto, na sondagem da satisfação do usuário foram coletadas opiniões de três categorias de usuários, não de duas. Deste modo, o teste *F ANOVA fator único* permitiu verificar se a hipótese nula deveria ou não ser aceita. A interpretação a ser dada aos casos nos quais a hipótese nula foi rejeitada deve limitar-se apenas à constatação de diferenças entre as médias, uma vez que não se pode afirmar que $\mu_P \neq \mu_I \neq \mu_E$.

Do mesmo modo que se procedeu para os dados coletados a partir da mensuração do desempenho do usuário, também foi necessária nesta análise a aplicação do procedimento *Tukey-Kramer*, a fim de determinar as médias significativamente diferentes dentre grupos de médias comparados.

A Tabela 9 ilustra os resultados obtidos com os testes *F ANOVA fator único* com $\alpha = 0,01$, realizados sobre os escores de cada seção do *OpUS*. Conforme pode ser observado a partir dos resultados dos testes administrados às três seções do *OpUS*, as estatísticas de teste *F* são suficientemente superiores aos F_{crit} determinados ou, alternativamente, que os *P-values* são bastante inferiores ao valor de α , de modo que a hipótese nula pode ser rejeitada para as três seções testadas.

Tabela 9 – Resultados da aplicação do teste *F* ANOVA fator único às três seções do OpUS.ANOVA (Single Factor) - Variável: **Escore da Seção 1** $\alpha = 0,01$

SUMMARY

Groups	Count	Sum	Average	Variance
PRINCIPIANTES	20	687	34,35	47,186842
INTERMEDIÁRIOS	20	693	34,65	52,239474
EXPERIENTES	20	358	17,90	7,989474

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	3675,0333	2	1837,5167	51,319736	1,79E-13	4,9981281
Within Groups	2040,9000	57	35,8053			
Total	5715,9333	59				

F > F_{crit} e P-value << α , então: **REJEITA-SE H₀**ANOVA (Single Factor) - Variável: **Escore da Seção 2** $\alpha = 0,01$

SUMMARY

Groups	Count	Sum	Average	Variance
PRINCIPIANTES	10	291	29,10	52,988889
INTERMEDIÁRIOS	10	374	37,40	34,933333
EXPERIENTES	10	182	18,20	6,844444

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	1854,4667	2	927,2333	29,35314808	1,68943E-07	5,488118404
Within Groups	852,9000	27	31,5889			
Total	2707,3667	29				

F > F_{crit} e P-value << α , então: **REJEITA-SE H₀**ANOVA (Single Factor) - Variável: **Escore da Seção 3** $\alpha = 0,01$

SUMMARY

Groups	Count	Sum	Average	Variance
PRINCIPIANTES	20	687	34,35	47,186842
INTERMEDIÁRIOS	20	693	34,65	52,239474
EXPERIENTES	20	358	17,90	7,989474

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	3675,0333	2	1837,5167	51,319736	1,79E-13	4,9981281
Within Groups	2040,9000	57	35,8053			
Total	5715,9333	59				

F > F_{crit} e P-value << α , então: **REJEITA-SE H₀**

Conforme anteriormente mencionado, o teste *F* ANOVA fator único foi complementado com a execução do procedimento *Tukey-Kramer*. A Tabela 10 ilustra os resultados do procedimento *Tukey-Kramer* com $\alpha = 0,01$ sobre os resultados do teste *F* ANOVA fator único.

Tabela 10 – Resultados da aplicação do procedimento Tukey-Kramer às três seções do OpUS.




COMPARAÇÕES MULTIPLAS DE TUKEY-KRAMER			
Seção do OpUS	$\alpha = 0,01$	Variável Analisada	
		Escore	
1	Média do Grupo 1	34,35	
	n do Grupo 1	20	
	Média do Grupo 2	34,65	
	n do Grupo 2	20	
	Média do Grupo 3	17,90	
	n do Grupo 3	20	
	MQD	35,8053	
	Estatística Q	4,157	
	COMPARAÇÃO ENTRE PRINCIPANTES E INTERMEDIÁRIOS		
	Diferença Absoluta	0,30	
	Erro Padrão da Diferença	1,338007159	
	Intervalo Crítico	5,562095761	
	As médias dos Grupos 1 e 2 são	Não Diferentes	
	COMPARAÇÃO ENTRE PRINCIPANTES E EXPERIENTES		
	Diferença Absoluta	16,45	
	Erro Padrão da Diferença	1,338007159	
	Intervalo Crítico	5,562095761	
	As médias dos Grupos 1 e 3 são	Diferentes	
	COMPARAÇÃO ENTRE INTERMEDIÁRIOS E EXPERIENTES		
	Diferença Absoluta	16,75	
Erro Padrão da Diferença	1,338007159		
Intervalo Crítico	5,562095761		
As médias dos Grupos 2 e 3 são	Diferentes		
2	Média do Grupo 1	29,10	
	n do Grupo 1	10	
	Média do Grupo 2	37,40	
	n do Grupo 2	10	
	Média do Grupo 3	18,20	
	n do Grupo 3	10	
	MQD	31,5889	
	Estatística Q	4,60	
	COMPARAÇÃO ENTRE PRINCIPANTES E INTERMEDIÁRIOS		
	Diferença Absoluta	8,30	
	Erro Padrão da Diferença	1,777326332	
	Intervalo Crítico	8,175701125	
	As médias dos Grupos 1 e 2 são	Diferentes	
	COMPARAÇÃO ENTRE PRINCIPANTES E EXPERIENTES		
	Diferença Absoluta	10,90	
	Erro Padrão da Diferença	1,777326332	
	Intervalo Crítico	8,175701125	
	As médias dos Grupos 1 e 3 são	Diferentes	
	COMPARAÇÃO ENTRE INTERMEDIÁRIOS E EXPERIENTES		
	Diferença Absoluta	19,20	
Erro Padrão da Diferença	1,777326332		
Intervalo Crítico	8,175701125		
As médias dos Grupos 2 e 3 são	Diferentes		
3	Média do Grupo 1	29,55	
	n do Grupo 1	20	
	Média do Grupo 2	32,25	
	n do Grupo 2	20	
	Média do Grupo 3	19,85	
	n do Grupo 3	20	
	MQD	87,0395	
	Estatística Q	4,157	
	COMPARAÇÃO ENTRE PRINCIPANTES E INTERMEDIÁRIOS		
	Diferença Absoluta	2,70	
	Erro Padrão da Diferença	2,086138462	
	Intervalo Crítico	8,672077588	
	As médias dos Grupos 1 e 2 são	Não Diferentes	
	COMPARAÇÃO ENTRE PRINCIPANTES E EXPERIENTES		
	Diferença Absoluta	9,70	
	Erro Padrão da Diferença	2,086138462	
	Intervalo Crítico	8,672077588	
	As médias dos Grupos 1 e 3 são	Diferentes	
	COMPARAÇÃO ENTRE INTERMEDIÁRIOS E EXPERIENTES		
	Diferença Absoluta	12,40	
Erro Padrão da Diferença	2,086138462		
Intervalo Crítico	8,672077588		
As médias dos Grupos 2 e 3 são	Diferentes		

Os resultados que aparecem destacados em azul na Tabela 10 são relativos às categorias de usuários para os quais as médias dos escores foram significativamente diferentes entre si. Analogamente, os resultados destacados em vermelho referem-se às categorias para as quais as médias dos escores não apresentaram diferenças estatisticamente significativas entre si. Por sua vez, o Quadro 50 ilustra graficamente as diferenças registradas *procedimento Tukey-Kramer* com $\alpha = 0,01$, por seção do *OpUS* e por categoria de respondentes, de acordo com os resultados quantitativos e qualitativos da Tabela 10.

No Quadro 50, similarmente ao Quadro 44 (seção 7.4.2), a leitura da síntese gráfica das diferenças entre pares de categorias de respondentes deverá ser feita horizontalmente, e.g., para a seção 1 do *OpUS* foram constatadas, através do *procedimento Tukey-Kramer*, diferenças estatísticas significativas apenas entre os pares de categorias *principiantes-experientes* e *intermediários-experientes*.

Quadro 50 – Síntese gráfica dos resultados do *procedimento Tukey-Kramer*, em termos das seções do *OpUS* e das categorias de respondentes.

CATEGORIAS DIFERENTES PROCEDIMENTO DE TUKEY-KRAMER		
Seção do <i>OpUS</i>	$\alpha = 0,001$	Variável Analisada <i>Escore</i>
1	P	E
	I	E
2	P	I
	I	E
3	P	E
	I	E

	Principiantes		Intermediários		Experientes
---	---------------	---	----------------	--	-------------

Como se pode verificar, a partir da Tabela 10 e/ou do Quadro 50, os resultados obtidos a partir da análise estatística dos escores relativos às diferentes categorias de respondentes, computados para as diferentes seções do *OpUS*, confirmam a análise preliminar do Quadro 49, feita na sub-seção precedente.

7.4.2.3 Problemas Identificados a partir da Sondagem da Satisfação do Usuário

Similarmente às sub-seções 7.2.1, 7.2.2, 7.2.3 e 7.4.3, esta sub-seção apresenta os problemas identificados no processo interativo do usuário com o produto avaliado, a partir da sondagem da satisfação do usuário no tocante ao produto. Alguns desses problemas também foram verificados tanto a partir da inspeção de conformidade do produto às *Partes 14, 15 e 16* do padrão internacional *ISO 9241* (vide sub-seções 7.2.1, 7.2.2 e 7.2.3) quanto através da mensuração do desempenho do usuário durante o uso do produto (vide sub-seção 7.4.3).

Do mesmo modo como se procedeu na sub-seção 7.4.3, serão apresentados e discutidos ao longo desta seção apenas os problemas cuja identificação não foi comum aos dois enfoques acima mencionados. Ao final desta sub-seção, será apresentado um sumário dos problemas e um parecer sobre o produto, com base no enfoque avaliatório tratado nesta seção.

Antes de iniciar o registro dos problemas identificados, vale a pena discutir alguns aspectos relativos à sondagem de características do usuário e de sua opinião sobre produtos ou serviços. A sondagem de características do usuário visa *a priori* a estimação de parâmetros (fatos populacionais) a partir (i) da coleta de informações sobre indivíduos que utilizam um dado produto ou serviço como suporte em atividades cotidianas ou (ii) do levantamento de práticas correntes de uma determinada comunidade usuária daquele produto ou serviço.

Deste modo, torna-se possível o delineamento do perfil do usuário típico de um dado produto ou serviço, tal como aquele apresentado ao final da sub-seção 7.5.1, cujo propósito, em nível da avaliação do produto ou serviço considerado, transcende o suporte à elaboração de roteiros de teste. Obviamente este será o propósito primordial da sondagem de características do usuário, pois fornecerá subsídios a diferentes encaminhamentos do processo de investigação do produto ou serviço a ser avaliado. Todavia, o delineamento do perfil do usuário também fundamentará os resultados da análise de sua opinião sobre o produto ou serviço avaliado: enquanto a sondagem da opinião do usuário responde o que, a sondagem de suas características fundamenta essas respostas ao evidenciar quem as formulou.

Outro aspecto da sondagem da opinião do usuário que merece ser comentado é o modo como os resultados obtidos contribuem para a identificação de "gargalos" no produto ou serviço avaliado. Diferentemente da mensuração do desempenho do usuário e da inspeção de conformidade a um padrão, para os quais os referenciais disponíveis¹⁰¹ possibilitam um delineamento mais nítido dos referidos "gargalos", a sondagem da opinião e, por extensão, da satisfação do usuário resulta usualmente em indicações menos diretas da natureza desses "gargalos".

Pode-se argumentar que a nitidez do delineamento de falhas em uma sondagem da opinião sobre qualquer tema é dependente do modo como o questionamento foi estruturado, bem como do detalhamento das questões formuladas. Tal argumento, embora pertinente, implica a elaboração de instrumentos de sondagem com um número maior de itens e/ou contendo questões formuladas com um grau de detalhamento tal que poderá desestimular os respondentes.

Eis porque os instrumentos correntes de sondagem de aspectos da usabilidade de produtos ou serviços não excedem o total de 50 itens, cada um dos quais formulados com um grau moderado de detalhamento, como ocorre, por exemplo, com o *DePerUSI* e o *OpUS*.

¹⁰¹ Indicadores quantitativos, para o caso da *mensuração do desempenho*; listas de princípios ergonômicos consagrados, para o caso da *inspeção de conformidade a padrões*.

Um terceiro aspecto pertinente ao contexto desta discussão diz respeito à sondagem da opinião do usuário sobre diferentes aspectos de produtos ou serviços como meio de aquisição de informações relativas à visão do usuário enquanto consumidor. Sob este prisma, o levantamento da opinião visa, direta ou indiretamente, sondar a satisfação que o uso do produto ou serviço confere ao consumidor e, por extensão, quão aceitável lhe é tal produto ou serviço.

Assim, uma síntese dos escores dos aspectos sondados de um produto ou serviço, tal como aquela apresentada nos Quadros 48 e 49, já pode satisfazer aos propósitos do desenvolvedor, sobretudo se os resultados forem favoráveis. Caso contrário, usualmente se adota um enfoque avaliatório cujos resultados possam complementar aqueles da sondagem da opinião do usuário do produto, a exemplo dos demais enfoques adotados nesta pesquisa.

O último aspecto aqui comentado, fortemente relacionado com o aspecto anterior, chama a atenção para a natureza dos resultados da sondagem da opinião e sua contribuição em processos avaliatórios de produtos ou serviços. Conforme discutido na seção 3.3, as técnicas de avaliação da usabilidade de produtos ou serviços possuem um caráter eminentemente *diagnóstico*, uma vez que sua orientação mais corrente aponta para a identificação, classificação e racionalização de *problemas*. Contudo, também é necessário levar em conta a discussão conduzida na seção 4.4, na qual foram confrontados os enfoques avaliatórios de natureza *quantitativa* e *qualitativa* (é aconselhável rever as informações do Quadro 15).

Assim, a integração de todas as considerações feitas nesta sub-seção às considerações feitas nas seções 3.3 e 4.4 conduzem à constatação de que, apesar de serem valiosas como indicadores de “gargalos” existentes em produtos ou serviços, as informações advindas de levantamentos da opinião do usuário não dão uma indicação tão detalhada desses “gargalos” quanto aquelas obtidas a partir de enfoques avaliatórios fundamentados quer na mensuração do desempenho do usuário, quer na inspeção de conformidade a padrões.

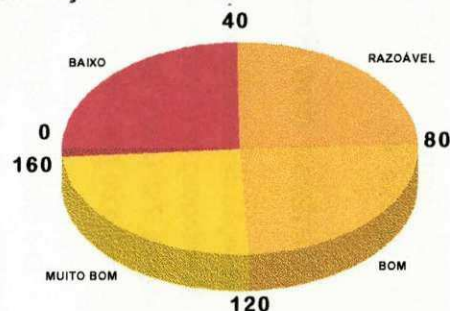
Por esta razão, a síntese dos resultados obtidos a partir da sondagem da opinião do usuário, conforme apresentado ao longo desta seção, terá um caráter distinto das sínteses apresentadas ao final das seções 7.2.3 (Quadro 40) e 7.4.4 (Quadro 46). Ao invés de um sumário de falhas, o Quadro 51 apresenta uma lista de indicações de aspectos do processo interativo usuário-produto que apresentaram problemas, sob o ponto de vista da opinião dos respondentes do *OpUS*.

Embutidas na listagem de problemas apresentada no Quadro 51, há uma série de falhas não identificáveis direta e precisamente a partir de levantamentos da opinião do usuário, i.e. a extensão do detalhamento das falhas será ampliada através da triangulação dos dados obtidos com dados resultantes da avaliação do produto ou serviço considerado a partir de outros enfoques. Tome-se como exemplo a confrontação de dados resultantes de um levantamento da opinião do usuário com dados provenientes da mensuração do desempenho do usuário e da inspeção de conformidade a um padrão (conforme será apresentada ao final deste capítulo).

Quadro 51 – Sumário de Problemas e Parecer sobre o **MATLAB v. 5.3.0** com base na Sondagem da Satisfação do Usuário.

# DO PROBLEMA	ITEM DO OpUS						COMENTÁRIO	ESCORE DO ITEM					
01	2						Revisar aspectos da comunicação usuário-produto, tais como terminologia, linguagem, retorno da informação e das ações em geral, a fim de otimizar o processo interativo.	74					
02	5						Revisar os mecanismos de acesso às instruções e advertências do produto, a fim de otimizar o processo interativo usuário-produto.	77					
03	8						Revisar a estruturação e a linguagem adotada nas instruções e advertências do produto, a fim de tornar mais fácil sua compreensão pelo usuário.	77					
04	11						Revisar as mensagens de erro apresentadas e redigi-las em estilo mais claro e menos lacônico.	56					
05	18						Assegurar a visualização à primeira vista das seqüências de ações necessárias para completar as tarefas de interesse com o auxílio do produto.	73					
06	26						Revisar a estruturação das informações contidas no manual do produto (ajuda <i>offline</i>) e alterá-la, a fim de tornar o processo de busca mais eficaz e efetivo.	57					
07	28						Analisar e melhorar aspectos do estilo de redação adotado no manual do produto (ajuda <i>offline</i>), tais como linguagem, clareza e conteúdo, a fim de facilitar a compreensão das informações oferecidas.	67					
08	30	38	42	43	44	45	Analisar a estruturação e o conteúdo da ajuda <i>online</i> e <i>offline</i> , assim como a eficácia de seus mecanismos, otimizando-os e/ou incrementando-os, conforme se faça necessário, a fim de (i) tornar mais fácil a superação de dúvidas e problemas encontrados pelo usuário durante o uso do produto; (ii) possibilitar a redução da freqüência das consultas durante o uso do produto; (iii) minimizar o tempo de consulta durante as sessões de trabalho; e (iv) elevar a eficiência do processo interativo.	75	79	38	59	46	66
09	36						Analisar os diferentes modos de interação com o produto (<i>menus</i> , <i>comandos</i> e <i>manipulação direta</i>), a fim de corrigir inconsistências que impeçam o usuário de executar de modo <u>direto</u> seqüências de ações com o auxílio do produto.	66					
10	41						Revisar a estruturação das funcionalidades oferecidas, a fim de facilitar o aprendizado do usuário.	49					
PARECER: A opinião do usuário sobre o MATLAB v. 5.3.0 resultou no escore final não ponderado 85,86 (correspondente ao escore final ponderado 53,66 , conforme apresentado nos Quadros 51 e 52) que, conforme a legenda abaixo apresentada, se faz equivaler à classificação BOM, indicando que a interface do produto <u>ainda</u> precisa ser analisada à luz dos problemas acima listados, a fim de que o processo de interação usuário-produto seja otimizado.													

CLASSIFICAÇÃO DO ESCORE INDIVIDUAL NÃO PONDERADO



A listagem do Quadro 51 se fundamentou no cruzamento das informações sumariadas nos Quadros 48 e 49, assim como em informações registradas em planilhas auxiliares, referentes às diferentes categorias de respondentes consideradas. A primeira coluna enumera os problemas apontados pelos respondentes do *OpUS*, enquanto a segunda explicita os itens do referido instrumento de sondagem que suscitaram os comentários feitos na terceira coluna do quadro.

A última coluna apresenta os escores brutos dos itens comentados do *OpUS*, assim como a classificação semântico-gráfica interpretável a partir da legenda anexa ao quadro. A última linha do Quadro 51 contém um parecer sobre o produto, com base no enfoque avaliatório adotado, similarmente ao que foi feito nos Quadros 40 e 46, respectivamente, para os enfoques fundamentados na inspeção de conformidade do produto às *Partes 14, 15 e 16* do padrão *ISO 9241* e na mensuração do desempenho do usuário durante o uso do produto.

É importante observar que as faixas *0-40, 41-80, 81-120 e 121-160* da legenda, associadas aos escores brutos dos itens do *OpUS*, correspondem às faixas ponderadas *0-25, 26-50, 51-75 e 76-100*. Assim, o escore global ponderado 53,66 corresponde à classificação semântica *BOM* da legenda, conforme mencionado no parecer apresentado ao final do Quadro 49.

7.5 Síntese dos Resultados Apresentados

A última seção deste capítulo apresenta, através dos Quadros 52 e 53, uma síntese dos resultados dos três enfoques avaliatórios adotados nesta pesquisa, os quais foram apresentados e discutidos nas seções 7.2 (*Inspeção de Conformidade com Padrão Internacional*), 7.4 (*Mensuração do Desempenho do Usuário*) e 7.5 (*Sondagem da Satisfação do Usuário*).

O Quadro 52 sumaria, ao mesmo tempo que confronta, as informações apresentadas nos Quadros 40, 46, 48 e 51, tendo sido construído com base na estrutura do *OpUS*. A 1ª coluna contém as declarações dos itens do instrumento de sondagem da opinião do usuário, em função das quais foi elaborada a confrontação dos resultados da sondagem da opinião do usuário com os demais enfoques avaliatórios adotados.

A 2ª coluna confronta os aspectos do processo interativo associados às declarações dos itens do instrumento de sondagem da opinião do usuário com os resultados obtidos a partir da inspeção de conformidade do produto às *Partes 14, 15 e 16* do padrão internacional *ISO 9241*.

Similarmente à 2ª coluna, as 3ª e 4ª colunas confrontam os aspectos do processo interativo associados às declarações dos itens do instrumento de sondagem da opinião do usuário com os resultados obtidos a partir da mensuração do desempenho do usuário durante o uso do produto e na sondagem da satisfação do usuário quanto ao produto avaliado, respectivamente.

Por fim, a última coluna do Quadro 52 confronta os resultados obtidos, respectivamente, a partir da inspeção de conformidade do produto às partes consideradas do padrão *ISO 9241*, da mensuração do desempenho do usuário e da sondagem da satisfação do usuário quanto ao produto, evidenciando similaridades e discrepâncias entre os três enfoques avaliatórios.

Quadro 52 – Confrontação dos resultados obtidos a partir dos diferentes enfoques avaliatórios adotados.

ITEM do OpUS	Resultados da Inspeção de Conformidade com o ISO 9241 PARTES 14, 15 e 16	Resultados da Mensuração do Desempenho do Usuário	Resultados da Sondagem da Satisfação do Usuário	COMENTÁRIOS
1. Uso do produto na realização de tarefas de interesse	Todas as falhas apontadas no Quadro 40 dificultam, em menor ou maior grau, o uso do produto, mesmo em tarefas relativamente simples conforme discutido nas sub-seções 7.2.1, 7.2.2 e 7.2.3	A Tabela 1 mostra que pelo menos 10,0% dos usuários de teste não conseguiram concluir as tarefas mais simples (1 e 3), enquanto pelo menos 57,5% não conseguiram concluir as tarefas mais complexas (2 e 4).	90,0% dos indivíduos se mostraram neutros (57,5%) ou classificaram o produto como de fácil uso (32,5%), os 10,0% restantes se dividiram igualmente entre as opções muito fácil e difícil.	Os resultados dos três enfoques avaliatórios adotados são convergentes, se a neutralidade registrada na análise da sondagem da satisfação for interpretada como indicação das categorias de respondentes principiantes e intermediários por insuficiência de familiaridade com o produto.
2. Comunicação com o produto (terminologia, linguagem, retorno da informação e das ações em geral)	As falhas 08 a 10, 16, 21, 22 e 26 apontadas no Quadro 40 dificultam, em menor ou maior grau, a comunicação com o produto.	As linhas relativas a ações incorretas da Tabela 1 (2ª linha associada a cada tarefa) mostra que para a tarefa considerada mais complexa (4) 72,5% dos usuários de teste cometeram pelo menos 3 ações incorretas em função da comunicação com o produto.	92,5% dos indivíduos classificaram a comunicação com o produto como difícil (35,0%), se mostraram neutros (32,5%) ou a classificaram como fácil (25,0%), os 7,5% restantes a classificaram como muito difícil (5,0%) ou muito fácil (2,5%).	Os resultados dos três enfoques avaliatórios adotados são convergentes.
3. Localização dos itens de menu associados às tarefas	O produto passou na inspeção de conformidade com a Parte 14 do ISO 9241 (diálogo via menus) no que diz respeito a este aspecto.	A Tabela 1 mostra que 70,0% dos usuários de teste não conseguiram concluir a tarefa mais complexa (4), enquanto 57,5% fizeram pelo menos 3 opções incorretas nesta tarefa, devido à localização dos itens do menu.	87,5% dos indivíduos se mostraram neutros (42,6%), classificaram o item como fácil (32,5%) ou muito difícil (12,5%), os 12,5% restantes classificaram-no como difícil (7,5%) ou muito fácil (5,0%).	Os resultados dos três enfoques avaliatórios adotados são divergentes, uma vez que os resultados da inspeção de conformidade da Parte 14 do ISO 9241 não condizem com os demais resultados obtidos.
4. Acesso aos itens de menu associados às tarefas	As falhas 03 e 04 apontadas no Quadro 40 dificultam o acesso aos itens de menu associados às tarefas.	A falha 18 do Quadro 49 explicita inconsistências no projeto do menu do produto que dificulta o acesso a opções desejadas e confunde o usuário.	92,5% dos indivíduos se mostraram neutros (45,0%), classificaram o acesso aos itens de menu difícil (32,5%) ou fácil (15,0%), os 7,5% restantes o classificaram como muito fácil.	Os resultados dos três enfoques avaliatórios adotados são convergentes.
5. Acesso às instruções e advertências do produto	As falhas 18 e 19 apontadas no Quadro 40 dificultam, em menor ou maior grau, o acesso às instruções e advertências do produto.	As falhas 01, 09, 10 e 11 do Quadro 49 dificultam, em menor ou maior grau, o acesso às instruções e advertências do produto.	37,5% dos indivíduos se mostraram neutros, 35,0% classificaram o acesso às instruções e advertências como difícil, 27,5% o classificaram como fácil.	Os resultados dos três enfoques avaliatórios adotados são convergentes.
6. Compreensão das instruções e advertências do produto	O produto passou na inspeção de conformidade com as Partes 14, 15 e 16 do ISO 9241 no que diz respeito a este aspecto. Todavia, é importante ressaltar que as declarações referentes a este aspecto não cobrem adequadamente.	62,1% dos usuários de teste apresentaram dificuldades na compreensão de termos e expressões, por conta da versão em inglês da interface. Além disso, 71,6% expressaram, através de comentários, dificuldades de compreensão de algumas mensagens do produto.	57,5% dos indivíduos se mostraram neutros; 25,0% classificaram a compreensão das instruções e advertências como difícil; 17,5% a classificaram como fácil.	Os resultados dos três enfoques avaliatórios adotados são divergentes, uma vez que os resultados da inspeção de conformidade do produto às Partes 14, 15 e 16 do ISO 9241 não condizem com os demais resultados obtidos.
7. Acesso às janelas de diálogo do produto	O produto passou na inspeção de conformidade com as Partes 14, 15 e 16 do ISO 9241 no que diz respeito a este aspecto.	Não foram identificadas falhas relacionadas com o acesso às janelas de diálogo, embora as falhas 03 e 14 do Quadro 49 estejam indiretamente relacionadas com o acesso às janelas de diálogo do produto.	87,5% dos indivíduos classificaram o acesso às janelas de diálogo como fácil (45,0%), se mostraram neutros (32,5%) ou classificaram-no como muito fácil (10,0%), os 12,5% restantes o classificaram como difícil (7,5%) ou muito difícil (5,0%).	Os resultados dos três enfoques avaliatórios adotados são convergentes.
8. Compreensão dos termos e solicitações apresentadas nas janelas de diálogo do produto	O produto passou na inspeção de conformidade com as Partes 14, 15 e 16 do ISO 9241 no que diz respeito a este aspecto.	Vide comentários feitos para os itens 5 e 6.	95,0% dos indivíduos classificaram a compreensão de termos e solicitações como fácil (35,0%), se mostraram neutros (32,5%) ou classificaram-na como difícil (27,5%), os 5,0% restantes se dividiram igualmente entre as opções muito fácil e muito difícil.	Os resultados da inspeção de conformidade e da sondagem da satisfação são convergentes, embora diverjam dos resultados da mensuração do desempenho. Todavia, é importante ressaltar que as declarações referentes a este aspecto (Partes 14, 15 e 16 do ISO 9241). Além disso, apesar de 67,5% dos indivíduos terem havido classificado este aspecto como fácil ou se mostrando neutros, 71,6% expressaram, através de comentários, dificuldades de compreensão de algumas mensagens do produto.
9. Recuperação de situações de erro	As falhas 11 a 17 apontadas no Quadro 40 dificultam, em menor ou maior grau, a recuperação de situações de erro.	Segundo a Tabela 1, 77,5% dos usuários de teste repetiram os mesmos erros pelo menos 4 vezes durante a execução da Tarefa 2, enquanto 50,0% o fizeram pelo menos 3 vezes durante a execução da Tarefa 4. Associado a estes fatos, 55,0% e 70,0% dos usuários de teste não conseguiram concluir, respectivamente, as Tarefas 2 e 4, devido, em grande parte, à repetição de erros.	92,5% dos indivíduos se mostraram neutros (45,0%), classificaram a recuperação de situações de erro como fácil (30,0%) ou difícil (17,5%), os 7,5% restantes a classificaram como muito difícil (5,0%) ou muito fácil (2,5%).	Os resultados dos três enfoques avaliatórios adotados são convergentes, se a neutralidade registrada na análise da sondagem da satisfação for interpretada como indicação das categorias de respondentes principiantes e intermediários por insuficiência de familiaridade com o produto.
10. Recuperação de situações de travamento	Este aspecto não foi contemplado nas declarações contidas nas Partes 14, 15 e 16 do ISO 9241.	20% dos usuários da teste fizeram comentários sobre situações de travamento durante a execução de scripts complexos. Todavia, o teste de usabilidade não registrou falhas relacionadas com este aspecto da interação.	75,0% dos indivíduos classificaram a recuperação de situações de travamento como difícil (32,5%), se mostraram neutros (30,0%) ou classificaram-na como fácil (12,5%), os 15,0% restantes a classificaram como muito difícil (12,5%) ou muito fácil (2,5%).	Não há dados suficientes para a confrontação dos enfoques adotados.
11. Compreensão das mensagens de erro apresentadas	A falha 17 apontada no Quadro 40 dificulta, em menor ou maior grau, a compreensão das mensagens de erro apresentadas. É conveniente ressaltar que as demais partes utilizadas do referido padrão não tratam suficientemente este aspecto.	Vide comentários feitos para os itens 5 e 6.	87,5% dos indivíduos se mostraram neutros (40,0%), classificaram a compreensão das mensagens de erro como fácil (25,0%) ou difícil (22,5%), os 7,5% restantes a classificaram como muito difícil (5,0%) ou muito fácil (2,5%).	Os resultados dos três enfoques avaliatórios adotados são convergentes, se a neutralidade registrada na análise da sondagem da satisfação for interpretada como indicação das categorias de respondentes principiantes e intermediários por insuficiência de familiaridade com o produto.
12. Navegação através das diferentes opções do menu, janelas de diálogo e barras de ícones do produto	As falhas 01 a 04, 10, 14 e 19 a 20 apontadas no Quadro 40 dificultam, em menor ou maior grau, a comunicação com o produto.	As falhas 01, 03, 07, 09, 12 e 15 e 18 a 20 do Quadro 49 implicam, em menor ou maior grau, dificuldades para a navegação do usuário através da estrutura de menus, janelas de diálogo e barras de ícones. Adicionalmente, a Tabela 1 mostra que pelo menos 40,0% dos usuários de teste cometeram pelo menos 3 ações incorretas associadas ao processo de navegação através diálogos de linha de comando e/ou manipulação direta durante a execução das tarefas mais simples (1 e 3), ao passo que 60,0% cometeram pelo menos 4 ações incorretas associadas ao processo de navegação através diálogos de linha de comando e/ou manipulação direta durante a execução das tarefas mais complexas (2 e 4).	77,5% dos indivíduos classificaram a navegação pelas opções de menu, janelas de diálogo e barras de ícones como fácil (37,5%), se mostraram neutros (25,0%) ou classificaram-na como difícil (15,0%), os 22,5% restantes a classificaram como muito fácil.	Os resultados dos inspeção de conformidade e da mensuração do desempenho são convergentes, embora diverjam dos resultados da sondagem da satisfação.
13. Navegação através de diferentes partes das tarefas executadas pelo produto	Vide comentário feito para o item 01.	Vide comentário feito para o item 12.	85,0% dos indivíduos se mostraram neutros (45,0%), classificaram a navegação através de diferentes partes das tarefas como fácil (25,0%) ou difícil (15,0%), os 15,0% restantes a classificaram como muito difícil (12,5%) ou muito fácil (2,5%).	Os resultados da inspeção de conformidade e da mensuração do desempenho são convergentes, embora diverjam dos resultados da sondagem da satisfação.
14. Memorização das seqüências de ações associadas a cada tarefa executada com o auxílio do produto	Nenhuma das declarações contidas nas Partes 14, 15 e 16 do ISO 9241 trata deste aspecto.	O teste de usabilidade realizado com o produto não contemplou este aspecto, nem foram registrados comentários sobre a memorização das seqüências de ações.	75,0% dos indivíduos se mostraram neutros (32,5%), classificaram este item como muito fácil (30,0%) ou difícil (12,5%), os 25,0% restantes o classificaram como fácil (22,5%) ou muito difícil (2,5%).	Não resultados a confrontar.
15. Aprendizado de novas seqüências de ações	Nenhuma das declarações contidas nas Partes 14, 15 e 16 do ISO 9241 trata do aprendizado de novas seqüências de ações pelo usuário.	Segundo a Tabela 1, 70,0% dos usuários de teste não conseguiram concluir a Tarefa 4, a qual envolvia novas seqüências de ações para 85,0% dos participantes dos testes, conforme registro dos comentários verbais. Além disso, 60,0% cometeram pelo menos 4 ações incorretas e 32,5% fizeram pelo menos 4 opções incorretas durante a execução da referida tarefa.	87,5% dos indivíduos classificaram o aprendizado de novas seqüências de ações como fácil (47,5%), se mostraram neutros (30,0%) ou classificaram-no como difícil (10,0%), os 12,5% restantes o classificaram como muito fácil (10,0%) ou muito difícil (2,5%).	Os resultados da mensuração do desempenho e da sondagem da satisfação são divergentes, conforme ilustram os resultados ao lado.
16. Uso das funcionalidades mais comuns do produto	As falhas apontadas no Quadro 40 dificultam até mesmo o uso das funcionalidades mais comuns do produto, conforme discutido nas sub-seções 7.2.1, 7.2.2 e 7.2.3.	Vide comentário feito para o item 01.	47,5% dos indivíduos classificaram o uso das funcionalidades mais comuns como fácil; 32,5% se mostraram neutros; 20,0% o classificaram como muito fácil.	Os resultados da mensuração do desempenho e da inspeção de conformidade são divergentes dos resultados obtidos a partir da sondagem da satisfação, uma vez que o posicionamento predominantemente favorável dos respondentes do OpUS contraz os resultados dos demais enfoques.
17. Exploração de novas funcionalidades do produto	Nenhuma das declarações contidas nas Partes 14, 15 e 16 do ISO 9241 trata explicitamente deste aspecto.	Vide comentário feito para o item 15.	95,0% dos indivíduos se mostraram neutros (37,5%), classificaram a exploração de novas funcionalidades como fácil (30,0%) ou difícil (27,5%), os 5,0% restantes se dividiram igualmente entre as opções muito fácil e muito difícil.	Os resultados da mensuração do desempenho e da sondagem da satisfação são convergentes apenas a neutralidade de 37,5% dos respondentes do OpUS for interpretada como indicação por insuficiência de familiaridade com o produto.

Quadro 52 – Confrontação dos resultados obtidos a partir dos diferentes enfoques avaliatórios adotados (continuação).

Item do OpUS	Resultados da Inspeção de Conformidade com o ISO 9241 Partes 14, 15 e 16	Resultados da Mensuração do Desempenho do Usuário	Resultados da Sondagem da Satisfação do Usuário	COMENTÁRIOS
18. Visualização à primeira vista das seqüências de ações necessárias para completar as tarefas de interesse com o auxílio do produto	Nenhuma das declarações contidas nas Partes 14, 15 e 16 do ISO 9241 trata explicitamente deste aspecto.	Durante a realização das tarefas pelos usuários, nas 40 sessões de teste foram registradas em 67,5% das sessões pelo menos 3 ações de "passivo aleatório" com o mouse por sobre as opções do menu do produto avaliadas em buscas não direcionadas e incertas por funcionalidades de interesse.	85,0% dos indivíduos classificaram o item como difícil (40,0%), se mostraram neutros (32,5%), classificaram-no como fácil (12,5%), os 15,0% restantes o classificaram como muito difícil.	Os resultados da mensuração do desempenho e da sondagem da satisfação são convergentes uma vez que denotam a existência de dificuldades na visualização à primeira vista das seqüências de ações necessárias para completar as tarefas de interesse com o auxílio do produto.
19. Visualização à primeira vista da estruturação dos menus, barras de ícones ou listas de informações disponibilizadas pelo produto	Nenhuma das declarações contidas nas Partes 14, 15 e 16 do ISO 9241 trata da exploração de novas funcionalidades do produto pelo usuário.	Vide comentário feito para o item 18.	92,5% dos indivíduos se mostraram neutros (40,0%), classificaram o item como fácil (32,5%) ou difícil (20%), os 7,5% restantes o classificaram como muito fácil (5,0%) ou muito difícil (2,5%).	Os resultados da mensuração do desempenho e da sondagem de satisfação são convergentes uma vez que denotam a existência de dificuldades na visualização à primeira vista da estruturação dos menus, barras de ícones ou listas de informações disponibilizadas pelo produto, por sua vez intimamente relacionada com a visualização imediata das seqüências de ações necessárias para completar as tarefas de interesse com o auxílio do produto.
20. Processo de entrada e saída de dados durante o uso do produto	As falhas 05 a 15 e 18 e 26 apontadas no Quadro 40 dificultam em menor ou maior grau o processo de entrada e saída de dados durante o uso do produto.	Praticamente todas as falhas apontadas no Quadro 49 apresentam uma relação direta ou indireta com este aspecto, em especial as falhas 04 a 12.	85,0% dos indivíduos classificaram o processo de entrada e saída de dados como difícil (47,5%), se mostraram neutros (25,0%) ou classificaram-no como fácil (12,5%), os 15,0% restantes o classificaram como muito fácil (10,0%) ou muito difícil (5,0%).	Os resultados dos três enfoques adotados são convergentes, uma vez que se complementam para traduzir a existência de problemas no processo de entrada e saída de dados.
21. Localização e acesso aos mecanismos de ajuda online do produto	As falhas 01 e 19 apontadas no Quadro 40 dificultam em menor ou maior grau a localização e acesso aos mecanismos de ajuda online do produto.	As falhas 01 e 15 apontadas no Quadro 49 dificultam em menor ou maior grau a localização e acesso aos mecanismos de ajuda online do produto.	72,5% dos indivíduos classificaram a localização e acesso à ajuda online como fácil (42,5%), difícil (17,5%) ou se mostraram neutros (12,5%), os 27,5% restantes o classificaram como muito fácil.	Os resultados da inspeção de conformidade e da mensuração do desempenho são convergentes, embora diverjam dos resultados da sondagem de satisfação.
22. Uso dos mecanismos de ajuda online do produto	As falhas 01 e 19 apontadas no Quadro 40 dificultam em menor ou maior grau o uso dos mecanismos de ajuda online do produto.	89,3% das consultas à ajuda online elou offline realizadas pelos usuários durante as 40 sessões de teste do produto não possibilitaram a superação de dúvidas e dificuldades. O comentário feito para o item 06 também se aplica a este item. Estas constatações refletem dificuldades no uso dos mecanismos de ajuda disponibilizados.	85,0% dos indivíduos classificaram o item como fácil (40,0%), difícil (27,5%) ou se mostraram neutros (17,5%), os 15,0% restantes o classificaram como muito fácil.	Os resultados da inspeção de conformidade e da mensuração do desempenho são convergentes, embora diverjam dos resultados da sondagem de satisfação.
23. Localização e acesso às informações de interesse na ajuda online	As falhas 01 e 19 apontadas no Quadro 40 dificultam em menor ou maior grau a localização e acesso às informações de interesse na ajuda online do produto.	Vide comentário feito para o item 22.	75% dos indivíduos se mostraram neutros (32,5%), classificaram a localização e acesso às informações de interesse na ajuda online como fácil (27,5%) ou difícil (15,0%), os 25,0% restantes a classificaram como muito difícil (10,0%), muito fácil (7,5%) ou se abstiveram de opinar (7,5%).	Os resultados da inspeção de conformidade e da mensuração do desempenho são convergentes, só convergindo com os resultados da sondagem de satisfação se a neutralidade dos 30% dos respondentes do OpUS for interpretada como insuficiência de familiaridade com o produto.
24. Compreensão das informações de interesse existentes na ajuda online	As declarações das Partes 14, 15 e 16 do ISO 9241 não contemplam explicitamente questões de clareza e eficácia das mensagens de erro e de ajuda.	Vide comentários feitos para os itens 06 e 22.	92,5% dos indivíduos se mostraram neutros (62,5%), classificaram a compreensão das informações de interesse existentes na ajuda online como fácil (17,5%) ou difícil (12,5%), os 7,5% restantes a classificaram como muito fácil.	Os resultados da mensuração do desempenho são convergentes com os da sondagem de satisfação se a neutralidade de 62,5% dos respondentes do OpUS for interpretada como insuficiência de familiaridade com o produto.
25. Linguagem utilizada na ajuda online	As declarações das Partes 14, 15 e 16 do ISO 9241 não contemplam explicitamente a linguagem adotada na ajuda.	68,5% dos usuários de teste declararam que a linguagem utilizada na ajuda online era clara e compreensível. No entanto, 89,3% das consultas à ajuda online elou offline realizadas pelos usuários durante as 40 sessões de teste do produto não possibilitaram a superação de dúvidas e dificuldades.	87,5% dos indivíduos se mostraram neutros (47,5%), classificaram a linguagem utilizada na ajuda online como difícil (22,5%) ou fácil (17,5%), os 12,5% restantes a classificaram como muito fácil.	Embora as constatações da observação das consultas à ajuda online contradigam os comentários verbais feitos por 68,5% dos usuários de teste, os resultados da observação e da sondagem da satisfação são convergentes.
26. Localização das informações de interesse no manual do produto (ajuda offline)	As declarações das Partes 14, 15 e 16 do ISO 9241 não tratam de questões relativas à ajuda offline.	Vide comentário feito para o item 22.	70% dos indivíduos classificaram localização das informações de interesse no manual como difícil (27,5%), fácil (25,0%) ou se mostraram neutros (17,5%), os 30,0% restantes a classificaram como muito difícil (15,0%), muito fácil (7,5%) ou se abstiveram de opinar (7,5%).	Os resultados da mensuração do desempenho e da sondagem de satisfação são divergentes, só convergindo se a neutralidade de 17,5% dos respondentes do OpUS, aliada à abstenção de 7,5% dos respondentes, forem interpretadas como insuficiência de familiaridade com o produto.
27. Uso do manual do produto (ajuda offline)	Vide comentário feito para o item 26.	Vide comentários feitos para os itens 22 e 25.	77,5% dos indivíduos se mostraram neutros (35,0%), classificaram o uso do manual como fácil (30,0%) ou difícil (12,5%), os 22,5% restantes se abstiveram de opinar (12,5%), classificaram o referido uso como muito fácil (7,5%) ou muito difícil (2,5%).	Os resultados da mensuração do desempenho e da sondagem de satisfação são divergentes.
28. Compreensão das informações de interesse existentes no manual do produto (ajuda offline)	Vide comentário feito para o item 26.	Vide comentário feito para o item 06.	85,0% dos indivíduos se mostraram neutros (37,5%), classificaram o item como fácil (27,5%) ou difícil (20,0%), os 15,0% restantes se abstiveram de opinar (12,5%) ou classificaram-no como muito fácil (2,5%).	Os resultados da mensuração do desempenho e da sondagem de satisfação são divergentes.
29. Linguagem utilizada no manual do produto (ajuda offline)	Vide comentário feito para o item 26.	Vide comentário feito para o item 25.	80% dos indivíduos classificaram a linguagem do manual como fácil (37,5%), se mostraram neutros (35,0%) ou classificaram-na como difícil (7,5%), os 20,0% restantes se abstiveram de opinar (12,5%) ou classificaram o item como muito fácil (7,5%).	Vide comentário feito para o item 25.
30. Superação de dúvidas e problemas encontrados durante as sessões de uso do produto a partir dos mecanismos oferecidos tanto pela ajuda online quanto pelo manual do produto (ajuda offline)	Vide comentário feito para o item 26.	89,3% das consultas à ajuda online elou offline realizadas pelos usuários durante as 40 sessões de teste do produto não possibilitaram a superação de dúvidas e dificuldades. O comentário feito para o item 06 também se aplica a este item.	82,5% dos indivíduos se mostraram neutros (52,5%) ou se dividiram igualmente entre as opções fácil (15,0%) e difícil (15,0%), os 17,5% restantes se abstiveram de opinar (10,0%) ou classificaram o item como muito difícil (7,5%).	Os resultados da mensuração do desempenho e da sondagem de satisfação são divergentes, só convergindo se a neutralidade de 52,5% dos respondentes do OpUS, aliada à abstenção de 10,0% dos respondentes, forem interpretadas como insuficiência de familiaridade com o produto.
31. Acho que a apresentação do produto é bastante atraente, o que estimula seu uso por mim.	As declarações das Partes 14, 15 e 16 do ISO 9241 não tratam deste aspecto.	De um modo geral, os comentários feitos pelos usuários da versão 5.2 do produto foram favoráveis às inovações feitas na apresentação da versão avaliada (5.3). O mesmo ocorreu com os usuários que já estavam utilizando a versão avaliada.	87,5% dos indivíduos concordaram parcialmente (47,5%), concordaram totalmente (27,5%) ou discordaram totalmente (12,5%) desta afirmativa, os 12,5% restantes discordaram parcialmente (7,5%) ou se mantiveram neutros (5,0%).	Os resultados da mensuração do desempenho e da sondagem de satisfação são convergentes.
32. Acho que a apresentação do produto facilita o aprendizado rápido de seus comandos.	Vide comentário feito para o item 31.	Vide comentário feito para o item 15.	67,5% dos indivíduos concordaram (45,0%) ou discordaram (22,5%) parcialmente desta afirmativa, dos 32,5% restantes, 30,0% se dividiram igualmente em discordância e concordância total com a afirmativa, enquanto 2,5% se mantiveram neutros.	Os resultados da mensuração do desempenho e da sondagem de satisfação são divergentes.
33. Acho o modo de apresentação das mensagens do produto bastante claro e compreensível.	Vide comentário feito para o item 24.	Vide comentários feitos para os itens 5 e 6.	85,0% dos indivíduos concordaram (32,5%) ou discordaram (27,5%) parcialmente desta afirmativa ou mantiveram-se neutros (32,5%), os 15,0% restantes dividiram-se igualmente em discordância e concordância total com a afirmativa.	Os resultados da mensuração do desempenho e da sondagem de satisfação são divergentes, só convergindo se a neutralidade de 32,5% dos respondentes do OpUS for interpretada como insuficiência de familiaridade com o produto.
34. A resposta do produto às minhas entradas é muito lenta.	O produto passou na inspeção das declarações das Partes 14, 15 e 16 do ISO 9241 referentes à rapidez das respostas.	Apenas 17,5% dos usuários de teste comentaram que o produto se mostrou lento às suas entradas acrescentando que o fato ocorreu via de regra quando este processo scripts mais complexos.	80,0% dos indivíduos discordaram totalmente (40,0%), concordaram parcialmente (22,5%) ou totalmente (17,5%) com esta afirmativa, os 20,0% restantes discordaram parcialmente (15,0%) ou se mantiveram neutros (5,0%).	Os resultados da mensuração do desempenho e da sondagem de satisfação são convergentes, embora diverjam daqueles obtidos a partir da inspeção de conformidade.
35. Sempre me sinto no controle das ações quando uso o produto.	Nenhuma das declarações contidas nas Partes 14, 15 e 16 do ISO 9241 trata do controle das ações pelo usuário.	Durante as 40 sessões de teste do produto, 72,5% dos participantes perderam o controle das ações pelo menos 2 vezes.	82,5% dos indivíduos dividiram-se igualmente em concordância, discordância ou neutralidade com relação a esta afirmativa, os 17,5% restantes dividiram-se em concordância (12,5%) e discordância (5,0%) total desta afirmativa.	Os resultados da mensuração do desempenho e da sondagem de satisfação são divergentes.

Quadro 52 – Confrontação dos resultados obtidos a partir dos diferentes enfoques avaliatórios adotados (conclusão).

ITEM do OpUS	RESULTADOS DA INSPEÇÃO DE CONFORMIDADE COM O ISO 9241 PARTES 14, 15 E 16	RESULTADOS DA MENSURAÇÃO DO DESEMPENHO DO USUÁRIO	RESULTADOS DA SONDAÇÃO DA SATISFAÇÃO DO USUÁRIO	COMENTÁRIOS
36. Consigo executar as tarefas de modo <u>direto</u> ao usar o produto.	Item não pertinente ao contexto da inspeção de conformidade do produto às Partes 14, 15 e 16 do ISO 9241	As falhas 03, 14, 15 e 17 a 20 do Quadro 49 são as que contrariam mais frontalmente, em diferentes estágios do processo interativo, a execução de tarefas de modo <u>direto</u> com o auxílio do produto.	75,0% dos indivíduos <u>discordaram</u> (30,0%) ou <u>concordaram</u> (27,5%) <u>parcialmente</u> desta afirmativa ou se mantiveram <u>neutros</u> (17,5%); os 25,0% restantes dividiram-se igualmente em <u>discordância</u> e <u>concordância total</u> com a afirmativa.	Os resultados da mensuração do desempenho e da sondagem da satisfação são <u>divergentes</u> .
37. Acho que o produto atende plenamente às minhas necessidades.	Item não pertinente ao contexto da inspeção de conformidade do produto às Partes 14, 15 e 16 do ISO 9241.	52,5% dos usuários de teste mencionaram conseguir fazer <u>exatamente</u> o que intentavam durante as sessões de uso do produto. O teste de usabilidade em ambiente controlado não possibilitou a exploração mais aprofundada deste item.	87,5% dos indivíduos <u>concordaram totalmente</u> (35,0%), <u>concordaram parcialmente</u> (32,5%) ou <u>discordaram parcialmente</u> (20,0%) com esta afirmativa; os 12,5% restantes <u>discordaram totalmente</u> (7,5%) ou mantiveram-se <u>neutros</u> (5,0%).	De acordo com os resultados disponíveis da mensuração do desempenho e da sondagem da satisfação são <u>convergentes</u> .
38. Perco muito tempo tentando aprender os comandos do produto.	As declarações das Partes 14, 15 e 16 do ISO 9241 não tratam do tempo de aprendizado dos comandos do produto.	Vide comentários feitos para os itens 1, 9, 12 e 15, assim como falhas 03, 04, 05, 08, 10, 11 e 12 do Quadro 49.	82,5% dos indivíduos <u>concordaram totalmente</u> (40,0%), <u>concordaram parcialmente</u> (27,5%) ou <u>discordaram parcialmente</u> (15,0%) com esta afirmativa; os 17,5% restantes mantiveram-se <u>neutros</u> (10,0%) ou <u>discordaram totalmente</u> (7,5%).	Os resultados da mensuração do desempenho e da sondagem da satisfação são <u>convergentes</u> .
39. Consigo fazer <u>exatamente</u> o que desejo com os recursos oferecidos pelo produto.	Item não pertinente ao contexto da inspeção de conformidade do produto às Partes 14, 15 e 16 do ISO 9241.	Vide comentário feito para o item 37.	87,5% dos indivíduos <u>concordaram parcialmente</u> (52,5%), <u>concordaram totalmente</u> (20,0%) ou <u>concordaram totalmente</u> (15,0%) com esta afirmativa; os 12,5% restantes mantiveram-se <u>neutros</u> (7,5%) ou <u>discordaram totalmente</u> (5,0%) desta afirmativa.	De acordo com os dados disponíveis, os resultados da mensuração do desempenho e da sondagem da satisfação são <u>convergentes</u> .
40. As funcionalidades oferecidas pelo produto são facilmente configuráveis de modo a atenderem às minhas necessidades.	As falhas 05 e 07 apontadas no Quadro 40 dificultam, em menor ou maior grau, a configuração das funcionalidades oferecidas pelo produto.	Durante a execução da Pré-Tarefa (0) e da Tarefa 1 (mais simples) que incluíram ações de configuração do caminho da pasta de trabalho do usuário, 42,5% e 57,5% dos usuários de teste cometeram, respectivamente, pelo menos 1 ação incorreta relacionada com as ações de configuração mencionadas.	85,0% dos indivíduos <u>concordaram parcialmente</u> (40,0%), mantiveram-se <u>neutros</u> (30,0%) ou <u>discordaram parcialmente</u> (15,0%) desta afirmativa; os 15,0% restantes dividiram-se em <u>discordância</u> (10,0%) e <u>concordância</u> (5,0%) total com a afirmativa.	Os resultados da inspeção de conformidade e da mensuração do desempenho são <u>convergentes</u> , embora <u>divirjam</u> dos resultados da sondagem da satisfação.
41. O produto apresenta tantas funcionalidades que às vezes acho que nunca aprenderei ou terei necessidade de usá-las todas.	As declarações das Partes 14, 15 e 16 do ISO 9241 não tratam da quantidade de funcionalidades oferecidas pelo produto.	Vide comentário feito para o item 18.	77,5% dos indivíduos <u>concordaram totalmente</u> (37,5%), <u>concordaram parcialmente</u> (25,0%) ou mantiveram-se <u>neutros</u> (15,0%) com relação a esta afirmativa; os 22,5% restantes <u>discordaram totalmente</u> (12,5%) ou <u>parcialmente</u> (10,0%) com a afirmativa.	Os resultados da mensuração do desempenho e da sondagem da satisfação são <u>convergentes</u> .
42. Ao usar o produto, consulto freqüentemente os mecanismos de ajuda oferecidos.	Item não pertinente ao contexto da inspeção de conformidade do produto às Partes 14, 15 e 16 do ISO 9241.	92,5% dos usuários de teste afirmaram consultar a <u>ajuda online</u> e/ou <u>offline</u> pelo menos 2 vezes durante cada sessão de uso do produto.	70,0% dos indivíduos <u>concordaram totalmente</u> (52,5%) ou mantiveram-se <u>neutros</u> (17,5%) com relação a esta afirmativa; os 30,0% restantes dividiram-se igualmente em <u>discordância</u> e <u>concordância parcial</u> com a afirmativa.	Os resultados da mensuração do desempenho e da sondagem da satisfação são <u>convergentes</u> .
43. Não acho que as informações da ajuda deste produto são suficientemente eficazes para tirar minhas dúvidas.	Vide comentário feito para o item 24.	77,5% dos usuários de teste comentaram sobre dificuldades enfrentadas na busca de informações de interesse na <u>ajuda online</u> e/ou <u>offline</u> .	50,0% dos indivíduos dividiram-se igualmente em <u>concordância total</u> ou <u>parcial</u> com esta afirmativa, enquanto 45,0% dividiram-se igualmente em <u>discordância parcial</u> ou <u>neutralidade</u> com relação à afirmativa; os 5,0% restantes dividiram-se igualmente em <u>discordância total</u> da afirmativa e <u>abstenção</u> da opinião.	Os resultados da mensuração do desempenho e da sondagem da satisfação são <u>convergentes</u> .
44. Ao consultar a ajuda do produto, sempre perco muito tempo tentando encontrar as informações de interesse.	As declarações das Partes 14, 15 e 16 do ISO 9241 não tratam de questões relativas à eficiência da ajuda do produto.	Vide comentário feito para o item 43.	85,0% dos indivíduos <u>concordaram totalmente</u> (45,0%), <u>concordaram parcialmente</u> (22,5%) ou <u>discordaram parcialmente</u> (17,5%) desta afirmativa; os 15,0% restantes mantiveram-se <u>neutros</u> (10,0%) ou <u>discordaram totalmente</u> (5,0%).	Os resultados da mensuração do desempenho e da sondagem da satisfação são <u>convergentes</u> .
45. De um modo geral, a quantidade de informações oferecidas pela ajuda do produto é <u>insuficiente</u> para solucionar meus problemas e dúvidas.	Item não pertinente ao contexto da inspeção de conformidade do produto às Partes 14, 15 e 16 do ISO 9241.	Não foram registrados comentários dos usuários de teste sobre este aspecto.	25,0% dos indivíduos <u>discordaram totalmente</u> desta afirmativa; 40,0% dividiram-se igualmente em <u>concordância</u> ou <u>discordância parcial</u> da afirmativa; 17,5% mantiveram-se <u>neutros</u> ; 12,5% <u>discordaram totalmente</u> da afirmativa; os 5,0% restantes se abstiveram de opinar.	Não há dados suficientes para a confrontação dos enfoques adotados.
46. De um modo geral, a qualidade das informações oferecidas pela ajuda do produto <u>não</u> contribui para a solução dos meus problemas e dúvidas.	Vide comentários feitos para os itens 24 e 26.	Não foram registrados comentários dos usuários de teste sobre este aspecto. No entanto, o comentário feito no item 30 tem uma relação estreita com este item.	87,5% dos indivíduos <u>discordaram totalmente</u> (35,0%), <u>concordaram parcialmente</u> (30,0%) ou <u>discordaram parcialmente</u> (22,5%) desta afirmativa; os 12,5% restantes <u>discordaram totalmente</u> (7,5%) da afirmativa ou mantiveram-se <u>neutros</u> (5,0%).	De acordo com os resultados disponíveis da mensuração do desempenho e da sondagem da satisfação são <u>divergentes</u> .
47. Sinto-me <u>receoso</u> em alguns momentos das sessões de uso do produto.	Item não pertinente ao contexto da inspeção de conformidade do produto às Partes 14, 15 e 16 do ISO 9241.	66,5% dos usuários de teste esboçaram pelo menos 1 reação de receio face a suas ações, durante a execução das Tarefas 2 e 4, as quais não conseguiram concluir.	75,0% dos indivíduos <u>concordaram parcialmente</u> (42,5%), <u>discordaram parcialmente</u> (17,5%) ou mantiveram-se <u>neutros</u> (15,0%) com relação a esta afirmativa; os 25,0% restantes dividiram-se em <u>concordância</u> (15,0%) e <u>discordância</u> (10,0%) total com a afirmativa.	O s resultados disponíveis da mensuração do desempenho e da sondagem da satisfação são <u>convergentes</u> .
48. Em alguns momentos, sinto-me <u>frustrado</u> com o modo como o produto executa tarefas de meu interesse.	Item não pertinente ao contexto da inspeção de conformidade do produto às Partes 14, 15 e 16 do ISO 9241.	57,5% dos usuários de teste esboçaram pelo menos 2 reações de frustração durante a execução das Tarefas 2 e 4, as quais não conseguiram concluir.	82,5% dos indivíduos dividiram-se igualmente em <u>concordância</u> , <u>discordância</u> ou <u>neutralidade</u> com relação a esta afirmativa; os 17,5% restantes dividiram-se em <u>concordância</u> (10,0%) e <u>discordância</u> (7,5%) total desta afirmativa.	Os resultados da mensuração do desempenho e da sondagem da satisfação são <u>divergentes</u> .
49. De um modo geral, sinto-me <u>satisfeito</u> ao usar o produto.	Item não pertinente ao contexto da inspeção de conformidade do produto às Partes 14, 15 e 16 do ISO 9241.	Não foram registrados comentários sobre este aspecto ao longo das sessões de teste	92,5% dos indivíduos <u>concordaram parcialmente</u> (42,5%), <u>concordaram totalmente</u> (37,5%) ou <u>discordaram parcialmente</u> (12,5%) desta afirmativa; os 12,5% restantes mantiveram-se <u>neutros</u> (7,5%) ou <u>discordaram totalmente</u> (2,5%) da afirmativa.	Não há dados suficientes para a confrontação dos enfoques adotados.
50. Recomendaria <u>sem hesitação</u> o uso do produto aos meus colegas.	Item não pertinente ao contexto da inspeção de conformidade do produto às Partes 14, 15 e 16 do ISO 9241.	A partir dos resultados do delineamento do perfil dos usuários de teste através do DePerUSI, o avaliador questionou cada participante dos testes sobre a indicação do produto a outros usuários potenciais. 92,5% dos participantes responderam afirmativamente a este questionamento informal.	90,0% dos indivíduos <u>concordaram totalmente</u> (52,5%), <u>concordaram parcialmente</u> (25,0%) ou <u>discordaram parcialmente</u> (12,5%) desta afirmativa; os 10,0% restantes mantiveram-se <u>neutros</u> .	Os resultados da mensuração do desempenho e da sondagem da satisfação são <u>convergentes</u> .

Por outro lado, o Quadro 53, que também sumaria as informações apresentadas nos Quadros 40, 46 e 51, apresenta a lista de problemas de usabilidade identificados a partir dos enfoques avaliatórios adotados nesta pesquisa, em nível genérico (resultados da análise das informações coletadas a partir da sondagem da satisfação do usuário quanto ao produto) e específico (resultados das análises das informações coletadas a partir da inspeção de conformidade do produto com *Partes 14, 15 e 16* do padrão internacional *ISO 9241* e mensuração do desempenho do usuário durante o uso do produto).

A 1ª coluna contém as declarações dos itens do instrumento de sondagem da opinião do usuário, em função das quais foi elaborada a confrontação dos resultados da sondagem da opinião do usuário com os demais enfoques avaliatórios adotados. É importante salientar que a numeração das falhas não corresponde àquela utilizada nos Quadros 40, 46 e 51, servindo apenas para referenciar, conforme necessário, falhas comentadas no texto.

A 2ª coluna, baseada nos resultados da sondagem da satisfação do usuário, apresenta os problemas de usabilidade em um nível mais genérico, conforme discutido na sub-seção anterior. Por sua vez, a 3ª coluna associa às células da 2ª coluna as falhas identificadas a partir da inspeção de conformidade e da mensuração do desempenho do usuário. Como se sabe, as falhas identificadas a partir destes últimos enfoques possuem um caráter mais específico do que aquelas identificadas a partir da sondagem da satisfação do usuário.

As células em cinza correspondem aos problemas de usabilidade em nível genérico para os quais não foram identificados problemas em nível específico a partir da inspeção de conformidade do produto às *Partes 14, 15 e 16* do padrão internacional *ISO 9241* e mensuração do desempenho do usuário durante o uso do produto.

Por fim, as três últimas colunas do Quadro 52, agrupadas sob o rótulo **Enfoque Adotado**, associam as falhas listadas na 2ª e 3ª colunas aos enfoques avaliatórios considerados nesta pesquisa, com o propósito de evidenciar a extensão de cada um deles no que diz respeito à identificação de falhas.

É conveniente observar que as células da última coluna - **SS – Sondagem da Satisfação** - foram dimensionadas em consonância com as células da 2ª coluna, indicando o caráter genérico do referido enfoque avaliatório no tocante à identificação de falhas. Por outro lado, as células das colunas **IC (Inspeção de Conformidade)** e **MD (Mensuração do Desempenho)** seguem o dimensionamento das células da 3ª coluna, tendo sido preenchidas em consonância com a lista de falhas explicitadas na referida coluna.

Conforme se observa no Quadro 53, a sondagem da satisfação do usuário no tocante ao produto, a partir dos 50 itens *OpUS*, permitiu a identificação de 10 problemas de usabilidade, em nível genérico, alguns dos quais (problemas 4, 6 e 9) não foram identificados em nível específico a partir dos demais enfoques de avaliação adotados. Aos demais problemas genéricos, foi identificada pelo menos uma falha em nível específico.

# DA FALHA	NÍVEL GENÉRICO	NÍVEL ESPECÍFICO	ENFOQUE ADOTADO		
			IC	MD	SS
01	Mecanismos de acesso às instruções e advertências do produto.	O ciclo de opções do menu principal é interrompido na transição da última opção (Help) para a primeira (File) e vice-versa.	/		
		Não há nenhum mecanismo que torne óbvio para o usuário o uso de teclas de funções e de atalhos no diálogo via linha de comando.	/	/	
		Não é permitida a redigitação ou edição apenas de trechos incorretos de comandos e parâmetros associados.	/	/	
		O sistema não interpreta nem aceita a entrada de comandos contendo erros de digitação e/ou ortografia, mesmo que não haja ambigüidade quanto ao comando pretendido.	/	/	
		Não há mecanismos para o processamento de comandos destrutivos (e.g., delete , clear).	/	/	
		Não há mecanismos para a reversão de ações do usuário (undo).	/	/	
		O sistema não oferece nenhuma realimentação sobre o tempo restante para a conclusão de processos em andamento.	/	/	
		Não há mecanismo de acesso à barra de ferramentas de manipulação direta via teclado (dispositivo de entrada alternativo do mouse)	/	/	/
		A barra de ferramentas de manipulações diretas do Editor/Debugger não possui ícone de retrocesso de ações (undo), o que impede o retorno do usuário a estados anteriores.	/	/	
		O sistema não permite o dimensionamento horizontal ou vertical de objetos gráficos, apenas o redimensionamento horizontal e vertical simultâneo de objetos "depositados" na janela do modelo.	/		
		O sistema não oferece mecanismos para a ampliação ou redução do conteúdo das janelas de resultados para fins de impressão.	/		
		O sistema não oferece mecanismos para a manipulação direta do fator de escala que, suave e continuamente, indiquem a alteração das dimensões de objetos gráficos.	/		
		O sistema não oferece mecanismos de movimentorolamento automático da tela quando do rearranjo de objetos gráficos na janela de modelos do Simulink .	/		
A rotação tridimensional é crítica , i.e. o mecanismo de ajuste dos atributos azimute e elevação é pouco eficiente , implicando tempo e esforço extras dos usuários.		/			
02	Estruturação e linguagem das instruções e advertências do produto.	Os designadores numéricos da lista de janelas abertas (opção Window do menu principal) são iniciados por 0 , não por 1 .	/		
		Não são oferecidos grupos alternativos de nomes de comandos como opções para diferentes categorias de usuários.	/		
		Várias palavras de comando excedem o limite de 7 caracteres estabelecido no padrão (há palavras com até 19 caracteres).	/	/	
		Não é permitido o uso de abreviações de nomes de comandos durante a digitação de frases de comando.	/		
		Não foi adotada uma regra simples para a abreviação dos nomes de comandos.	/		
		Também não foi adotado o truncamento dos nomes de comandos, mas um misto de nomes completos e truncamento (e.g., cls sys proptable pagesetuplg).	/		
		O sistema executa qualquer sentença de comando correta sem solicitar antes nenhuma confirmação do usuário.	/	/	
		Os botões Home e Tips nas janelas de ajuda Help Windows e Help Tips contrariam a seqüência de manipulação direta de objetos, estando habilitados em circunstâncias em que não o deveriam.	/	/	/
		Não há indicação visual de onde um objeto gráfico do Simulink pode ser "depositado" pelo usuário no processo de criação de um modelo (arquivo .mdl).	/		
		Não há diferenciação semântica no cursor durante o processo de arrasto de um objeto gráfico da janela do Simulink para a janela do modelo, o que pode confundir o usuário.	/		
		A linguagem de comandos utiliza alguns nomes e sintaxes de comandos similares àquelas utilizadas pelo MS-DOS , porém não mantém a consistência em todas as circunstâncias em que se constata tal similaridade, o que confunde o usuário e propicia o cometimento de erros.		/	
		O projeto do painel de menu da opção Edit da janela de figuras é inconsistente , apresentando sub-opções em estados (habilitadas ou desabilitadas) que contrariam tanto os contextos de uso quanto os estados das sub-opções de outros painéis do menu principal da janela.		/	
		A barra de ferramentas de manipulações diretas da janela de figuras não possui ícone de retrocesso de ações (undo), o que impede o retorno do usuário a estados anteriores.		/	
03	Estruturação e linguagem das mensagens de erro.	O sistema não evidencia os erros cometidos em linhas de comando (as mensagens de erro são lacônicas e pouco esclarecedoras).	/	/	/
		O sistema não explicita a natureza (e.g. semântica, sintática, de orientação) do erro cometido, nem destaca visualmente as mensagens de erro das frases de comando, o que, em diversas situações, desnoiteia o usuário e dificulta sua recuperação face aos erros cometidos.		/	
04	Seqüências de ações necessárias para completar as tarefas de interesse com o auxílio do produto			/	
06	Estruturação e linguagem adotadas na redação do manual do produto (ajuda <i>offline</i>).			/	
07	Mecanismos de ajuda <i>online</i> e <i>offline</i> e conteúdo das informações contidas na ajuda.	Não há distinção entre objetos não selecionáveis (títulos e textos explicativos) e objetos selecionáveis (ítems consultáveis) nas modalidades de ajuda Help Windows e Help Tips	/	/	/
08	Execução de seqüências de ações de modo direto com o auxílio do produto.	O prompt da linha de comando não apresenta o caminho corrente, forçando o usuário a usar os comandos pwd e dir para "se situar" na estrutura de diretórios (dispêndio de tempo extra).		/	
		Antes de cada tentativa de reprocessamento de scripts e/ou de dados corrigidos, é necessário o uso do comando clear para a eliminação de resíduos de dados de processamentos anteriores no espaço de trabalho. Este requisito não é óbvio , forçando usualmente o usuário a "quebrar" a seqüência direta de ações, o que propicia o cometimento de erros de processamento de dados e pode confundir o usuário.		/	
		O retrocesso de ações (undo) executadas sobre objetos localizados na janela de gráficos, a partir da sub-opção Undo do menu (painel da opção Edit da janela de figuras, não é possível , uma vez que a sub-opção permanece desabilitada independentemente das ações executadas na área do gráfico (construção de linhas, alteração de propriedades dos eixos, de objetos gráficos ou de objetos de texto), o que usualmente "quebra" a seqüência direta de ações do usuário.		/	/
		A barra de ferramentas de manipulações diretas da janela de figuras não possui ícone de retrocesso de ações (undo), o que impede o retorno do usuário a estados anteriores.		/	
09	Estruturação e mecanismos da ajuda <i>online</i> e <i>offline</i> .			/	
10	Estruturação das funcionalidades oferecidas.	Não há nenhuma tecla de atalho para mover o cursor entre grupos de opções (e.g. entre o grupo Undo e o grupo Cut/Copy/Paste/Clear da opção Edit do menu principal).		/	/

Transcendendo a apresentação das séries de informações coletadas a partir dos três enfoques avaliatórios considerados e das análises isoladas das referidas informações, o presente capítulo também apresentou os resultados da *triangulação* das informações coletadas, i.e., a confrontação dos resultados obtidos a partir das técnicas de avaliação adotadas. Neste contexto da avaliação da usabilidade de produtos, a triangulação de dados objetiva a otimização do procedimento avaliatório, a partir da identificação de similaridades e discrepâncias nos resultados de diferentes enfoques considerados. Apesar de ser uma idéia comum, adotada por diversos pesquisadores (e.g. Jeffries *et al.* [Jeff91], Karat *et al.* [Kara92], Nielsen e Philips [Niel93d], Doubleday *et al.* [Doub97]), a triangulação envolve claramente o risco de que as recomendações e conclusões dos avaliadores, advindas da confrontação dos dados resultantes dos enfoques confrontados, possam ser pouco eficazes ou inócuas para os projetistas.

Evidentemente, ao se adotar um dado enfoque avaliatório ou a integração de um conjunto de enfoques, é necessário que se faça uma reflexão sobre as diversas condições restritivas¹⁰² oferecidas pelo contexto no qual se insere o foco da avaliação. Também é necessário que se proceda um julgamento das técnicas de avaliação associadas ao enfoque avaliatório considerado. Enfim, é necessário que se analise a metodologia que se pretende adotar à luz de um conjunto de critérios de aplicabilidade ao contexto e de confiabilidade dos resultados esperados, conforme propuseram Queiroz e Turnell [Quei98a, Quei98b].

Neste sentido, o mérito deste capítulo não reside na averiguação de qual dos enfoques adotados é o melhor, até porque tal averiguação implicaria esforços não triviais, senão sobre-humanos. O mérito reside na possibilidade de se estabelecer paralelos deste contexto com outros similares, a partir (i) da apresentação detalhada dos resultados obtidos a partir de cada um dos referidos enfoques; (ii) do detalhamento das análises não estatísticas e estatísticas efetuadas e (iii) da confrontação dos resultados com base nas limitações de cada enfoque.

O próximo capítulo apresentará as conclusões advindas desta pesquisa, comentando a aplicabilidade e abrangência dessas conclusões e propondo, através de novos tópicos de estudo, a continuidade desta linha de pesquisa.

¹⁰² E.g. recursos infra-estruturais e humanos disponíveis, disponibilidade de usuários de teste, custos associados, prazos a cumprir, abrangência do enfoque avaliatório.

8

CONSIDERAÇÕES FINAIS

"The human-centric technologies will bring computers closer to us and give us power to do more by doing less. But the highest meaning of "human-centric," and its biggest to us, will be determined by what we do to achieve the human goals we set. We will better off and we will be finishing the ultimate Unfinished Revolution if we reach for these goals using all our human dimensions in concert, standing once again in awe before the sunset, the wheel, and what may lie behind them."

[Michael L. Dertouzos , *The Unfinished Revolution* (2001)]

8	Considerações Finais	303
8.1	Conclusões	304
8.2	Aplicabilidade dos Resultados da Pesquisa	309
8.3	Proposições para Trabalhos Futuros	311

Encerrando o presente documento, este capítulo apresenta as conclusões advindas das análises dos resultados obtidos a partir dos três enfoques sobre os quais se respaldaram as hipóteses formuladas no início da pesquisa e apresentadas no primeiro capítulo deste documento.

A seção introdutória – **Conclusões** – resgata as hipóteses que nortearam o desenvolvimento do trabalho, apresentando as conclusões formuladas segundo a análise dos resultados obtidos mediante os enfoques avaliatórios considerados.

Na seção 8.2 (**Aplicabilidade e Extensão dos Resultados da Pesquisa**), são discutidos alguns aspectos relativos à aplicabilidade dos resultados obtidos, assim como à abrangência das inferências que resultaram nas conclusões formuladas na seção 8.1. Conforme enfatizado ao longo deste documento, a discussão conduzida nesta seção se fundamenta nos recursos materiais e humanos envolvidos no contexto da pesquisa, no universo amostral considerado e na metodologia adotada.

Por fim, a última seção deste capítulo propõe linhas de continuidade e/ou de aprofundamento do objeto desta pesquisa, sob a forma de sugestões, algumas das quais já se encontram em estruturação, como propostas de pesquisa, no âmbito do *Laboratório de Interfaces Homem-Máquina (LIHM)* do DEE/CCT/UFPB.

8.1 Conclusões

Conforme explicitado na seção 1.4, o desenvolvimento desta pesquisa intentou investigar a veracidade das seguintes afirmações:

- H1: A usabilidade será tanto maior quanto maior for o desempenho associado ao uso do produto.
- H2: A usabilidade será tanto maior quanto maior for a satisfação subjetiva associada ao uso do produto.
- H3: A usabilidade será tanto maior quanto maior for o grau de adequação do produto a um padrão internacional.
- H4: O desempenho do usuário durante o uso do produto será tanto maior quanto maior for a satisfação subjetiva associada ao uso do processo/produto.
- H5: A satisfação subjetiva será tanto maior quanto maior for o desempenho associado ao uso do processo/produto considerado.
- H6: O desempenho do usuário durante o uso do produto será tanto maior quanto mais fundamentado em um padrão (internacional) for o desenvolvimento do produto considerado.
- H7: A satisfação subjetiva será tanto maior quanto mais fundamentado em um padrão (internacional) for o desenvolvimento do produto considerado.

que, de acordo com a revisão de Estatística Indutiva apresentada no Anexo L representam as hipóteses nulas que nortearam tanto a revisão bibliográfica, apresentada nos Capítulos 3 e 4, quanto a aplicação dos procedimentos metodológicos descritos nos Capítulos 5 e 6.

A confrontação das hipóteses formuladas com as considerações resultantes da análise dos resultados obtidos (Capítulo 7), à luz do contexto avaliatório considerado, possibilitaram as seguintes conclusões:

Argumento 01: Se o número de falhas identificado nesta pesquisa a partir da mensuração do desempenho tivesse sido menor do que aquele apresentado no Quadro 49, os valores dos indicadores objetivos considerados¹⁰³ teriam sido mais baixos, o que implicaria um desempenho melhor do usuário e, por conseguinte, maior usabilidade do produto avaliado.

Conclusão 01: **A usabilidade será tanto maior quanto maior for o desempenho associado ao uso do produto.**
(Hipótese nula H1 *aceita*)

¹⁰³ Tempo de execução da tarefa, número de ações incorretas, número de opções incorretas, número de erros repetidos e número de consultas à ajuda.

Argumento 02: De acordo com os resultados obtidos a partir da sondagem da opinião do usuário, vários aspectos¹⁰⁴ do processo interativo do usuário com o produto avaliado receberam pontuações favoráveis e/ou corresponderam ao escore bom, quando computados, e.g. entrada e saída de dados, documentação *online* e *offline*, clareza das mensagens, facilidade de configuração das funcionalidades oferecidas, qualidade das informações da ajuda. No entanto, os referidos resultados não refletem a realidade associada aos ângulos de visão das demais estratégias de avaliação utilizadas nesta pesquisa, segundo os quais o processo interativo se mostrou, em maior ou menor grau, deficiente no que diz respeito aos aspectos mencionados.

Conclusão 02: **A usabilidade de um produto não será necessariamente maior quanto mais elevado for o nível de satisfação do usuário no tocante ao uso do produto.** (Hipótese nula H2 *rejeitada*)

Argumento 03: A conformidade de um produto às declarações contidas em um padrão internacional tendem a tornar o produto mais usável, uma vez que as declarações foram fundamentadas em axiomas e corolários que traduzem as iniciativas e esforços da comunidade pesquisadora para otimizar o processo interativo usuário-produto.

Conclusão 03: **A usabilidade de um produto será tanto maior quanto maior for o grau de adequação do referido produto às recomendações de um padrão internacional.** (Hipótese nula H3 *aceita*)

Argumento 04: O registro em vídeo das sessões de teste do produto facilitou a confrontação dos resultados da mensuração do desempenho e da sondagem da satisfação. Os indicadores objetivos de usabilidade mensurados, em associação com as informações subjetivas obtidas a partir da observação direta, por sua vez refinadas pelos resultados das análises retrospectivas das sessões registradas em vídeo, permitiram a identificação de várias discrepâncias entre os referidos enfoques avaliatórios (vide Quadro 55).

Conclusão 04: **Um nível mais elevado de satisfação subjetiva do usuário no tocante a um produto não implica necessariamente maior desempenho no uso do referido produto.** (Hipótese nula H5 *rejeitada*)

¹⁰⁴ E.g., itens 3, 8, 15, 16, 20, 24, 25, 33, 40 e 46 do OpUS.

Argumento 05: Conforme mencionado no argumento 3, alguns aspectos do processo interativo não foram contemplados nas recomendações das *Partes 14, 15 e 16* do padrão internacional *ISO 9241*, embora estivessem intimamente vinculados a sua usabilidade. Por outro lado, aproximadamente 60% das falhas identificadas a partir da mensuração do desempenho do usuário durante o uso do produto já haviam sido identificadas a partir da inspeção de conformidade do produto às *Partes 14, 15 e 16* do padrão internacional *ISO 9241*. Mesmo levando-se em conta que o documento-base da inspeção venha a excluir de seu teor aspectos do processo interativo relativos à usabilidade do produto avaliado, quanto maior for o número de recomendações aplicáveis adotadas no projeto, menor o número de falhas contidas no produto final, resultando em maior desempenho do usuário durante o uso do produto.

Conclusão 05: **O desempenho do usuário durante o uso do produto será tanto maior quanto mais fundamentado em um padrão (internacional) for o desenvolvimento do produto considerado.** (Hipótese nula H6 *aceita*)

Argumento 06: Conforme se observa no Quadro 55, identifica-se várias *discrepâncias*¹⁰⁵ entre os resultados obtidos a partir da inspeção de conformidade do produto às *Partes 14, 15 e 16* do padrão internacional *ISO 9241* e a mensuração da satisfação do usuário com relação ao produto. Além disto, vários aspectos considerados na sondagem da satisfação do usuário não foram contemplados pelas partes consideradas do padrão internacional adotado na inspeção de conformidade.

Conclusão 06: **Um nível mais elevado de satisfação subjetiva do usuário no tocante a um produto não implica necessariamente maior fundamentação do desenvolvimento daquele produto em um padrão (internacional).** (Hipótese nula H7 *rejeitada*)

Dando continuidade à série de conclusões apresentadas nesta seção, as análises dos resultados obtidos (Capítulo 7), à luz do contexto avaliatório considerado, possibilitaram as seguintes conclusões adicionais:

Argumento 07: A mensuração do desempenho do usuário, durante o uso do produto, possibilitou a identificação de problemas *de consistência e recorrentes superficiais e intermediários*, segundo a classificação apresentada no Quadro 48. Por outro lado, o Quadro 14 aponta a técnica de avaliação *Observação* como adequada à identificação de problemas *gerais e recorrentes graves*. Evidentemente, conforme discutido no capítulo anterior, é necessário ponderar esta informação, uma vez que se

¹⁰⁵ E.g., itens 8, 12, 13, 16, 21, 22 e 34 do *OpUS*.

demonstrou a adequação da referida técnica à identificação de outras classes de problemas.

Conclusão 07: **A mensuração do desempenho do usuário, durante o uso do produto, a partir da observação, também se adequa à identificação de problemas de consistência e recorrentes superficiais e intermediários.**

Argumento 08: A realização de análises retrospectivas de ações e reações dos usuários, durante as sessões de teste e ao longo do preenchimento do *OpUS*, deixou entrever que as reações neutras e positivas dos usuários a diferentes aspectos do processo interativo com o produto, corresponderam, em geral, a pontuações neutras e positivas aos itens equivalentes do *OpUS*. Por outro lado, verificou-se que a atribuição pontuações negativas a determinados aspectos do processo interativo sondados a partir do *OpUS* corresponderam a reações negativas dos usuários àqueles aspectos durante as sessões de teste.

Conclusão 08: **Partindo da premissa que a opinião do usuário reflete sua satisfação no que concerne ao uso de um produto, quanto melhor sua opinião sobre tal produto maior sua satisfação ao usá-lo.**

Argumento 09: Apesar dos padrões internacionais serem instrumentos importantes para a otimização da usabilidade de produtos, o uso de um padrão internacional merece a atenção do avaliador no que se refere à abrangência das declarações formuladas. Conforme se verificou no processo de inspeção de conformidade do produto avaliado às *Partes 14, 15 e 16* do padrão internacional *ISO 9241*, alguns aspectos do processo interativo não são considerados nas referidas partes, apesar de estarem intimamente vinculados à usabilidade de produtos de *software*. A segunda coluna do Quadro 55 atesta a veracidade desta afirmação.

Conclusão 09: **Independentemente do contexto avaliatório, o uso de padrões em processos de inspeção de conformidade deve ser complementado com uma avaliação das recomendações contidas no padrão utilizado, a fim de que se possa verificar sua abrangência e extensão e, por conseguinte, redobrar a atenção nos aspectos do processo interativo menos focalizados pelas recomendações.**

Conclusão 10: **A responsabilidade pela aprovação ou reprovação de um produto no que concerne às recomendações inclusas em um padrão recai muito mais sobre a interpretação dada pelo avaliador às declarações e às**

notas anexas contidas no padrão do que sobre a verificação pura e simples do produto quanto à adoção da recomendação.

Argumento 10: O registro em vídeo das sessões de teste do produto, incorporado ao processo de observação das diferentes categorias de usuários, durante a execução de tarefas, possibilitou a realização de análises retrospectivas de ações e reações dos usuários, facilitando a confrontação dos resultados da mensuração do desempenho e da sondagem da satisfação.

Conclusão 11: **O registro em vídeo de sessões de teste de usabilidade de produtos constitui uma ferramenta valiosa na identificação de discrepâncias entre estratégias de avaliação fundamentadas na mensuração do desempenho do usuário e em sua opinião sobre o produto usado.**

Argumento 11: A apresentação e discussão dos problemas de usabilidade identificados a partir da inspeção de conformidade do produto avaliado às *Partes 14, 15 e 16* do padrão internacional *ISO 9241* (Capítulo 7, sub-seções 7.2.1 a 7.2.3) mostrou que este enfoque avaliatório oferece subsídios para a identificação das causas dos problemas, i.e., as razões para os problemas identificados (e.g., o uso de títulos semanticamente indistintos em um menu) . Por outro lado, a apresentação e discussão dos problemas de usabilidade identificados a partir da mensuração do desempenho do usuário durante o uso do produto (Capítulo 7, sub-seção 7.4.3) mostrou que este enfoque possibilita a identificação dos “sintomas” dos problemas, ou seja, explicita os problemas em si, a partir dos erros observados (e.g., a escolha incorreta de uma opção de menu devido à similaridade semântica com outras opções).

Conclusão 12: **Apesar de compartilharem a mesma meta – a identificação de problemas de usabilidade em uma interface, a inspeção de conformidade de um produto a um padrão e os testes de usabilidade produzem resultados de naturezas sensivelmente diferentes.**

Conclusão 13: **O emprego da técnica de observação sem o respaldo de outra técnica que forneça informações sobre as causas os problemas observados (e.g., a inspeção de conformidade a um padrão) poderá ter implicações nas atividades de reprojeto, uma vez que a eliminação de um sintoma sem a eliminação de sua causa poderá implicar o surgimento posterior de novos sintomas.**

Argumento 12: A análise dos escores atribuídos por categoria de usuário aos itens de cada seção do *OpUS* permitiu constatar que os escores ponderados cresceram da categoria *principiantes* para a categoria *experientes*, exceto na Seção 2 do referido instrumento de sondagem, na qual o escore associado à categoria *intermediários* é ligeiramente superior ao da categoria *experientes*. Esta ordenação se mantém com relação aos escores globais ponderados das categorias consideradas.

Conclusão 14: **Quanto maiores o tempo e a frequência de uso de um produto, mais favoráveis tenderão a ser as opiniões dos respondentes sobre os aspectos daquele produto investigados a partir de um instrumento de sondagem da opinião do usuário.**

8.2 Aplicabilidade e Extensão dos Resultados da Pesquisa

Neste ponto, é importante analisar a aplicabilidade e a extensão dos resultados obtidos ao longo do desenvolvimento desta pesquisa. Antes de tudo, vale a pena lembrar que o foco da pesquisa não recaiu sobre o *MATLAB v. 5.3.0*, embora uma das contribuições deste trabalho pretenda ser as recomendações destinadas à otimização do processo interativo usuário-produto, advindas das limitações identificadas a partir de cada um dos enfoques de avaliação adotados. Entretanto, esta contribuição adviria como consequência da aplicação de um enfoque ou de um conjunto de enfoques avaliatórios, em um contexto qualquer de uso de um produto de *software*, independentemente do produto selecionado como alvo do processo avaliatório.

Na verdade, o foco da pesquisa foi direcionado para a confrontação de três enfoques avaliatórios centrados na usabilidade de produtos interativos usuário-computador – *mensuração do desempenho do usuário*, *mensuração da satisfação do usuário* e *inspeção de conformidade do produto a um padrão*, conforme explicitado no Capítulo 1 e reforçado ao longo dos demais capítulos deste documento, sendo o propósito da confrontação a investigação da influência do atributo central de cada enfoque considerado, i.e. o *desempenho*, a *satisfação subjetiva* e o *grau de conformidade a um padrão*, sobre a usabilidade da interface de um aplicativo de *software*.

Neste sentido, buscou-se desenvolver procedimentos metodológicos e instrumentos de suporte à aplicação de cada um dos enfoques avaliatórios considerados, visando não somente a realização da pesquisa, mas a incorporação de cada um deles aos contextos práticos do trabalho cotidiano realizado no *Laboratório de Interfaces Homem-Máquina (LIHM)* do DEE/CCT/UFPB. Conforme citado ao longo deste documento, à medida em que se fez pertinente, as iniciativas de Queiroz e Turnell [Quei96, Quei97a, Quei97b, Quei98a, Quei98b] resultaram na efetivação dos procedimentos metodológicos, detalhados nos Capítulos 5 e 6 desta tese, junto ao *LIHM/DEE/CCT/UFPB*.

Além da aplicação neste trabalho de pesquisa, os referidos procedimentos avaliatórios já haviam sido aplicados por Queiroz e Turnell [Quei98a] em um sistema de informações geográficas

(SIG) e, mais recentemente, por Nigam [Niga00] e Almeida [Alme00] a um banco de dados e SIG distribuídos na Web, respectivamente, tendo se mostrado adequado e levado a resultados satisfatórios em todos os casos. Estes estudos atestam a aplicabilidade e extensão dos procedimentos metodológicos descritos nos Capítulos 5 e 6 a diferentes contextos de uso de aplicações de *software*, assim como reflete a incorporação desses procedimentos às atividades de avaliação de interfaces usuário-computador realizadas no LIHM/DEE/CCT/UFPB.

No tocante à essência da pesquisa, uma análise retrospectiva dos estudos fundamentados nas diferentes técnicas empregadas na avaliação da usabilidade de interfaces mostra que apenas o trabalho de Doubleday *et al.* [Doub97], direcionado para a interface *INTUITIVE (Interactive User Interface Tools In a Visual Environment)* do projeto *ESPRIT*, apresenta similaridades com esta pesquisa. Conforme descrito na sub-seção 3.2.2.3, os autores compararam métodos de avaliação da usabilidade aplicados ao contexto da interface *INTUITIVE*, tendo confrontado os resultados de ensaios de usabilidade envolvendo observação direta, ensaio retrospectivo e uso de questionários com resultados obtidos a partir de sessões de avaliação heurística fundamentadas nas heurísticas propostas por Nielsen [Niel93b, Niel94a, Niel97c].

Embora parte do processamento dos dados coletados por Doubleday *et al.* se assemelhe ao processamento dos dados obtidos nesta pesquisa, os autores não exploraram em suas conclusões os resultados da sondagem da satisfação do usuário. Além disto, o universo amostral considerado pelos autores (metade do universo amostral considerado nesta pesquisa) foi tratado no âmbito da estatística de pequenas amostras, ao contrário desta pesquisa, tratada no âmbito das grandes amostras, o que restringe a extensão de suas inferências. Por fim, suas conclusões se limitam em mostrar a complementaridade dos resultados da avaliação heurística e da observação que respaldou seus testes de usabilidade.

Deste modo, resta mencionar, em nível das contribuições proporcionadas pelo desenvolvimento deste trabalho, que apesar das conclusões apresentadas na seção anterior serem restritas à aplicação selecionada como alvo de avaliação, assim como ao universo amostral de usuários da aplicação em nível do *Campus II* da UFPB, a extensão dessas conclusões transcende a dos autores acima citados.

Além dos procedimentos metodológicos, os demais resultados desta pesquisa também são aplicáveis a todas as linhas de ação do LIHM/DEE/CCT/UFPB, i.e. ensino, pesquisa e extensão, em especial à extensão. A metodologia descrita nos Capítulos 5 e 6, o material de suporte elaborado¹⁰⁶ (Capítulo 6), os procedimentos estatísticos e não estatísticos de análise dos dados coletados (Capítulo 7) e as conclusões registradas na seção anterior possibilitarão a programação de atividades de avaliação de produtos no âmbito do LIHM. Esta contribuição é relevante, principalmente para o LIHM, uma vez que viabiliza a prestação de serviços de avaliação, em nível da linha de ação *Extensão*, conforme a proposição feita por Queiroz e Turnell [Quei97c]. Também se insere nesta contribuição a infra-estrutura montada para a realização deste trabalho, a qual será, sem dúvida, aprimorada à medida que outros trabalhos de pesquisa incrementem as atividades atuais do LIHM/DEE/CCT/UFPB.

¹⁰⁶ E.g. fichas de registros de eventos, instrumentos de delineamento do perfil dos usuários de teste e de sondagem de sua opinião sobre os produtos sob condições de avaliação.

8.3 Proposições para Trabalhos Futuros

Ao longo dos Capítulos 5, 6 e, sobretudo, 7, foram explicitadas ou ficaram implícitas nas discussões dificuldades encontradas e questionamentos formulados durante o desenvolvimento desta pesquisa, alguns dos quais foram contornados para não entravar o andamento das atividades, embora ainda mereçam soluções mais eficazes.

Uma das dificuldades que merece ser investigada e mais adequadamente solucionada diz respeito ao registro dos eventos da observação associada à mensuração do desempenho do usuário durante o uso do produto sob condições de avaliação. Para esta pesquisa, foram revisadas e melhoradas as fichas de registro de eventos, também utilizadas por Nigam [Niga00] e Almeida [Alme00]. Mesmo assim, o registro manual de eventos ainda se mostrou ineficaz nas situações em que os eventos relativos à execução das tarefas de teste pelos usuários transcorreram em sucessões de ações excessivamente rápidas para serem observadas e registradas simultaneamente.

Previendo tais situações, incorporou-se ao processo de observação o registro em vídeo das sessões de teste. Apesar de ser um instrumento valioso em análises retrospectivas de seqüências de ações transcorridas em ritmos difíceis de observar e registrar simultaneamente, o registro em vídeo também exige uma carga elevada de concentração do observador, quando este atua simultaneamente como operador, como ocorre em laboratórios de usabilidade com uma equipe de avaliação reduzida, a exemplo do *LIHM/DEE/CCT/UFPB*. Face a esta dificuldade, sugere-se a automação do registro da observação, a partir do desenvolvimento de uma ferramenta que se adeque ao contexto de trabalho do *LIHM/DEE/CCT/UFPB*, similar às ferramentas *DRUM* [MacI93, MacI96], *The Observer* [NIT00] e *ErgoLight Lab Tester* [EL00], citadas no Capítulo 3.

Outro tópico que vale a pena ser explorado diz respeito ao refinamento dos instrumentos de sondagem elaborados nesta pesquisa, especialmente o *OpUS*. Conforme discutido no início da sub-seção 6.3.2, o *OpUS* originou-se do questionário *pós-teste* adaptado e utilizado por Queiroz [Quei94] na sondagem da satisfação de um grupo de usuários de um sistema de processamento de imagens multiespectrais. O referido questionário continha duas escalas de sondagem, uma para a avaliação do usuário sobre o item sondado, outra para sondagem da importância daquele item para o usuário. Esta última escala foi suprimida no *OpUS*, visando simplificar a atribuição da pontuação dos itens pelos respondentes, a exemplo de outros instrumentos de sondagem comercializados internacionalmente (e.g. a versão atualizada do *QUIS* [Harp97], o *ASQ* [Lewi93] e o *SUS* [Broo96]), conforme discutido na sub-seção 6.3.1. Similarmente a estes instrumentos de sondagem, a satisfação do usuário não foi mensurada neste trabalho, mas inferida da opinião global do usuário sobre o produto avaliado.

No entanto, de acordo com a revisão bibliográfica apresentada no Capítulo 3, sabe-se que este não é o único procedimento existente. Outros pesquisadores, e.g. Bailey e Pearson [Bail83], Ives *et al.* [Ives83] ou Estevam [Este90], computaram índices de satisfação do usuário a partir das pontuações atribuídas pelos respondentes aos aspectos sondados e à importância desses aspectos para seus contextos de trabalho. Portanto, se afigura válida a sugestão de incorporação da escala de importância dos itens ao *OpUS* e a administração deste instrumento sem e com escala de importância a uma amostra de usuários de um dado produto de *software*, com o

propósito de investigar a relevância da inclusão da referida escala. Vale salientar que a investigação sugerida possibilitaria adicionalmente a otimização do *OpUS*, tanto para fins de aplicação em nível do *LIHM/DEE/CCT/UFPB*, quanto para fins de comercialização.

Um segundo questionamento feito sobre os resultados da aplicação dos procedimentos metodológicos descritos nos Capítulos 5 e 6 aos contextos de uso considerados tanto por Queiroz e Turnell [Quei98a], Nigam [Niga00] e Almeida [Alme00], quanto nesta pesquisa, se refere ao modo de apresentação das falhas identificadas e dos pareceres fornecidos pelos avaliadores sobre os produtos avaliados. Embora os resultados do trabalho de Nigam [Niga00] tenham sido repassados para o desenvolvedor do *LightBase (Light Infocon)*, para fins de otimização do produto e apesar de se pretender repassar os resultados desta pesquisa para os desenvolvedores do *MATLAB (The MathWorks Inc.)*, com os mesmos fins, não houve um envolvimento efetivo entre as equipes de avaliação e de desenvolvimento dos referidos produtos, que facilitasse o uso de tais resultados. Tal "lacuna" deve ser investigada, a fim de que se possa avaliar a eficácia dos laudos técnicos elaborados pela equipe de avaliação sobre a manutenção do produto realizada pelos membros da equipe de desenvolvimento dos produtos avaliados. Sugere-se portanto, o desenvolvimento de um trabalho de pesquisa que envolva no contexto avaliatório do *LIHM/DEE/CCT/UFPB* um indivíduo atuante em uma equipe de desenvolvimento de *software*, integrando os pontos de vista do avaliador e do desenvolvedor, o que possibilitará o preenchimento da "lacuna" anteriormente mencionada.

A última sugestão para trabalho futuro advém da idéia de programação das atividades de avaliação de produtos do *LIHM/DEE/CCT/UFPB*, mencionada na seção anterior. Ao construir os Quadros 13 e 14 (Capítulo 4), Queiroz e Turnell [Quei96, Quei97b] elaboraram conjuntamente um fluxograma, com o propósito de facilitar a seleção de técnicas a serem adotadas na avaliação da interface de um sistema de processamento digital de imagens multiespectrais. Através dos vários níveis de decisão desse fluxograma, um grupo de técnicas de avaliação foi selecionado em função (i) do grau de detalhamento dos resultados esperados e (ii) dos recursos *infra-estruturais e humanos disponíveis*. Ao longo do desenvolvimento desta pesquisa, a viabilização da prestação de serviços de avaliação suscitou a idéia de estender o propósito do fluxograma mencionado à programação de atividades de avaliação de produtos no âmbito do *LIHM*, incluindo os fatores limitantes (iii) *prazos estipulados para a apresentação dos resultados* e (iv) *recursos econômicos disponíveis pelos clientes*. A generalização do referido fluxograma para um contexto específico produzirá um instrumento de suporte à seleção de técnicas de avaliação tanto para o *LIHM* quanto para outros laboratórios emergentes de avaliação de interfaces com o usuário.

A pesquisa científica não é estanque. Ao longo do desenvolvimento de trabalhos científicos, os pesquisadores se deparam com questões e dificuldades que buscam contornar. Todavia, muitas vezes algumas dessas questões e dificuldades se afiguram secundárias à linha central da investigação ou surgem ao longo do processamento e/ou das análises dos resultados coletados. Outras só são suscitadas após a ponderação das conclusões advindas ao final das análises. Antes de deixar no pesquisador a sensação de um trabalho incompleto, tais questões devem servir de fio condutor para o desenvolvimento de trabalhos de pesquisa futuros. Este trabalho de pesquisa não representa uma exceção à regra, razão pela qual se propôs nos parágrafos anteriores desta seção a série de tópicos de pesquisa que poderão evoluir fundamentados no mesmo eixo central da pesquisa que os originou.

REFERÊNCIAS BIBLIOGRÁFICAS

- [Aber90] ABERNETHY, C. N., Human-Computer Interface Standards: Origins, organizations and comment. In: D. J. Oborn (Ed.), **International Review in Ergonomics**, 2, 1990, pp. 31-54.
- [ABNT00a] ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, Institucional - Apresentação. **Documento Eletrônico**. www.abnt.org.br/instit_apresen.htm.
- [ABNT00b] ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, Institucional - Apresentação. **Documento Eletrônico**. www.abnt.org.br/instit_apresen.htm.
- [Abra77] ABRAMS, M. D. & TREU, S., A Methodology for Interactive Computer Service Measurement, **Communications of the ACM**, 20(12), December 1977, pp. 936-944.
- [Aike96] AIKEN, L. R., **Rating Scales & Checklists – Surveying Opinions and Assessing Personality**. John-Wiley & Sons, Inc., New York, 1996.
- [Aike97] AIKEN, L. R., **Questionnaires & Inventories – Evaluating Behavior, Personality, and Attitudes**. John-Wiley & Sons, Inc., New York, 1996.
- [Albe94] ALBEN, L.; FARIS, J. & SADDLER, H., Making It Macintosh: Designing the message when the message is design, **Interactions**, 1(1), January 1994, pp. 11-20.
- [Alle82] ALLEN, R. B., Cognitive Factors in Human Interaction. In: Albert Badre & Ben Shneiderman (Eds.), **Directions in Human-Computer Interaction**. Norwood, 1982, pp. 01 -26.
- [Alle84] ALLEN, R. B., Working paper on ethical issues for research on the use of computer services and interfaces, **SIGCHI Bulletin**, 16(1), July 1984, pp. 12-16.
- [Alme00] ALMEIDA, CECIR B. DE F., **Testes de Usabilidade para Planejamento de Sistemas Distribuídos de Informações Geográficas na Web**. (Dissertação de Mestrado em Informática) - DCC/CCT/UFPB, Campina Grande, Dezembro, 2000.
- [Amen99] AMENTO, B.; HILL, W.; TERVEEN, L.; HIX, D. & JU, P., An Empirical Evaluation of User Interfaces for Topic Management of Web Sites. CHI'99 Conference on Human Factors in Computing Systems: the CHI is the limit. **Proceedings**. Pittsburgh, PA, May 15 - 20 1999, pp. 552-559.
- [AMI92] AMI, **A quantitative approach to software management**, South Bank University for the AMI Consortium, ESPRIT AMI-5494, 1992.
- [Anas68] ANASTASI, A., **Psychological Testing**. Collier-Macmillan Ltd, London, 1968, 3rd edition.
- [App192] APPLE COMPUTER INC., **Apple - Human Interface Guidelines**. Apple Computer Inc., Cupertino - CA, 1992.
- [App197] APPLE COMPUTER, INC., **Mac OS 8 Human Interface Guidelines**. Apple Computer, Inc., Cupertino - CA, 1997.

- [Ayki94] AYKIN, NURAY, Reuse: A Case Study on Cost-Benefit of Adopting a Common Software Development Tool. In: R. G. Bias & D. J. Mayhew (Eds.), **Cost-Justifying Usability**. Academic Press, Chesnut Hill, Massachusetts, 1994, pp. 177-202.
- [Babe97] BABER, R. L., Comparison of electrical "engineering" of Heaviside's times and software "engineering" of our times, **IEEE Annals of the History of Computing**, 19(4), October-December 1997, pp. 5-17. (Versão online disponível em www.computer.org/annals/an1997/a4toc.htm).
- [Bach95] BACH, CLAUDIA, The Standards Process: Evolution or Revolution?, **StandardView**, 3(1), March 1995, pp. 29-32.
- [Badr99] BADRE, A. & JACOBS, A., Usability, Aesthetics, and Efficiency: An Evaluation in a Multimedia Environment. In: IEEE International Conference on Multimedia Computing and Systems ICMCS'99. Florence, Italy. 7 - 11 June 1999, Volume I, pp. 103-106.
- [Bail83] BAILEY, J. E. & PEARSON, S. W., A Tool for Computer User Satisfaction, **Management Science**, 29(5), May 1983, pp. 530-545.
- [Bail93] BAILEY, GREGG, Iterative Methodology and Designer Training in Human-Computer Interface Design. In: INTERCHI'93 - CHI'93 & INTERACT'93 Conference, 1993. **Proceedings**. Amsterdam, 1993, pp. 198-205.
- [Bain79] BAINBRIDGE, L., Verbal reports as evidence of process operator's knowledge, **International Journal of Man-Machine Studies**, 11, 1979, pp. 411-436.
- [Bala99] BALAKRISHNAN, R. & KURTENBACH, G., Exploring Bimanual Camera Control and Object Manipulation in 3D Graphics Interfaces. In: CHI'99 Conference on Human Factors in Computing Systems: the CHI is the limit. **Proceedings**. Pittsburgh, PA, May 15 - 20 1999, pp. 56-62.
- [Bala97] BALAKRISHNAN, R.; BAUDEL, T.; KURTENBACH, G. & FITZMAURICE, G., The Rockin'Mouse: Integral 3D Manipulation on a Plane. In: Conference on Human Factors in Computing Systems. **Proceedings**. Atlanta, GA, May 12 - 27 1997, pp. 311-318.
- [Ball92] BALLAS, J. A.; HEITMEYER, C. L. & PÉREZ, M., Evaluating two aspects of driver manipulation in advanced cockpits. In: CHI'92 Conference on Human Factors in Computing Systems. **Proceedings**. Monterey, CA, May 3 - 7 1992, pp. 127-134.
- [Bark89] BARKI, H. & HARTWICK, J., Rethinking the Concept of User Involvement, **MIS Quarterly**, 13(1), March 1989, pp. 53-63.
- [Baro86] BAROUDI, J. J., OLSON, M. H. & IVES, B., An Empirical Study of the Impact of User Involvement on System Usage and Information Satisfaction, **Communications of the ACM**, 29(3), March 1986, pp. 232-238.
- [Baro88] BAROUDI, J. J. & ORLIKOWSKI, W. J., A Short Form measure of User Information Satisfaction: A Psychometric Evaluation and Notes on Use, **Journal of Management Information Systems**, 4(4), 1988, pp. 45-59.
- [Baru97] BARUA, A.; CHELLAPA, R. & WHINSTON, A. B., Social Computing: Computer Supported Cooperative Work and Groupware. In: Gavriel Salvendy (Ed.), **Handbook of Human Factors and Ergonomics**, pp. 1760-1782. New York: John Wiley & Sons Inc. 1997.
- [Basi88] BASILI, V. R. & ROMBACH, H. D., The TAME Project: Towards Improvement-Oriented Software Environments, **IEEE Transactions on Software Engineering**, June 1988, pp. 758-773.

- [Basi94a] BASILI, V.R.; CALDIERA, G. & ROMBACH, H.D., GQM Paradigm. **Encyclopaedia of Software Engineering**, vol. 1, John Wiley & Sons, 1994, pp. 469-476.
- [Basi94b] BASILI, V.R.; CALDIERA, G. & ROMBACH, H.D., Experience Factory. **Encyclopedia of Software Engineering**, vol. 1, John Wiley&Sons, 1994, pp. 469-476.
- [Bass81] BASS, L. J. & BUNKER, R. E., A Generalized User Interface for Applications Programs, **Communications of the ACM**, 24(12), December 1981, pp. 796-800.
- [Bass85] BASS, L. J., A Generalized User Interface for Applications Programs, **Communications of the ACM**, 28(6), June 1985, pp. 617-627.
- [Baue72] BAUER, F. L., Software Engineering, **Information Processing**, 71, 1972.
- [Baum00] BAUMEISTER, L. K.; JOHN, B. E. & BYRNE, M. D., A comparison of tools for building GOMS models. In: CHI'2000 conference on Human factors in computing systems. **Proceedings**. The Hague Netherlands, April 1 - 6, 2000, pp. 502-509.
- [Bear96] BEARD, D. V.; SMITH, D. K. & DENELSBECK, K. M., Quick and Dirty GOMS: A Case Study of Computed Tomography, **Human-Computer Interaction**, 11 (2), 1996, pp.157-180.
- [Bell91] BELL, B.; RIEMAN, J. & LEWIS, C., Usability Testing of a Graphical Programming System: Things We Missed in a Programming Walkthrough. In: ACM CHI'91 Conference, 1991. **Proceedings**. New Orleans, Louisiana, April 27 - May 2, Louisiana, 1991, pp. 7-12.
- [Beme91] BEMELMANS, T. M. A., **Bestuurlijke Informatiesystemen en Automatisering**. Stenfert Kroese, The Netherlands, 1991.
- [Benb93] BENBASAT, I. & TODD, P., An experimental investigation of interface design alternatives: icons vs. text and direct manipulation vs. menus, **International Journal of Man-Machine Studies**, 38, pp. 369-402.
- [Benn84] BENNETT, J. L., Managing to meet usability requirements: establishing and meeting software development goals. In: J. Bennett, D. Case, J. Sandelin & M. Smith (Eds.), **Visual Display Terminals**. Prentice-Hall, Englewood Cliffs - NJ, 1984, pp. 161-184.
- [Benn86] BENNETT, J. L., Observations on Meeting Usability Goals for Software Products, **Behavioural Information Technology**, 5, pp. 183-193.
- [Berg90] BERG, J. L. & SCHUMMY, H., **An Analysis of the Information Technology Standardization Process**. Elsevier Science Publishers, Amsterdam, the Netherlands, 1990.
- [Berr88] BERRY, R. E., Common User Access - A Consistent and Usable Human-Computer Interface for the SAA Environment, **IBM Systems Journal**, 27(3), March 1988, pp. 281-300.
- [Berr90] BERRY, D. C. & BROADBENT, D. E., The role of instruction and verbalization in improving performance on complex search tasks, **Behaviour & Information Technology**, 9(3), May-June 1990, pp. 175-190.
- [Beva91] BEVAN, N.; KIRAKOWSKI, J. & MAISSEL, J., What is Usability? In: 4th International Conference on HCI. **Proceedings**. Stuttgart, September 1991. (Versão *online* disponível em info.lut.ac.uk/research/husat/inuse/usabilitypapers.html).
- [Beva93] BEVAN, N. & HOLDAWAY, K., User needs for user system interaction standards. In B. L. Meek; C. D. Evans & R. S. Walker (Eds), **User Needs in Information Technology Standards**, Computer Weekly Professional Series, chapter 5, pages 148-160. Butterworth-Heinemann, January 1993.

- [Beva94] BEVAN, N. & MACLEOD, M., Usability measurement in context, **Behaviour and Information Technology**, 13, 1994, pp.132-145. (Versão *online* disponível em info.lut.ac.uk/research/husat/inuse/usabilitypapers.html).
- [Beva95a] BEVAN, NIGEL, Measuring usability as quality of use, **Software Quality Journal**, 4, 1995, pp. 115-150 (Versão *online* disponível em info.lut.ac.uk/research/husat/inuse/usabilitypapers.html).
- [Beva95b] BEVAN, NIGEL, Usability is Quality of Use. In: 6th International Conference on Human-Computer Interaction, 1995. **Proceedings**. Yokohama, July 1995. Anzai & Ogawa (Eds.), Elsevier. (Versão *online* disponível em info.lut.ac.uk/research/husat/inuse/usabilitypapers.html).
- [Beva95c] BEVAN, NIGEL, Human-Computer Interaction Standards. In: 6th International Conference on Human-Computer Interaction, 1995. **Proceedings**. Yokohama, July 1995. Anzai & Ogawa (Eds.), Elsevier.
- [Beva97] BEVAN, NIGEL, Quality and usability: A new framework. In: van Veenendaal, E. & McMullan, J. (Eds.), **Achieving software product quality**. Tutein Nolthenius, Netherlands, 1997. (Versão *online* disponível em info.lut.ac.uk/research/husat/inuse/usabilitypapers.html).
- [Beva98] BEVAN, NIGEL, Usability Issues in Web Site Design Version 3. In: Usability Professional Association Conference UPA'98. **Proceedings**. Washigton DC, 22-26 June 1998. (Versão *online* disponível em www.usability.serco.com/pubs.html).
- [Beye94] BEYER, H. & HOLTZBLATT, K., Calling Down the Lightning, **IEEE Software**, 27(9), September 1994, pp. 106-107 e 113.
- [Bias91] BIAS, R., Walkthroughs: Efficient collaborative testing, **IEEE Software**, 8(5), September 1991, pp. 224-228.
- [Bias94] BIAS, R. & MAYHEW, D. J. (Eds.), **Cost-Justifying Usability**. Academic Press, Boston, 1994.
- [Bier92] BIERMAN, A. W.; FINEMAN, L. & HEIDLAGE, J. F., A voice- and touch-driven natural language editor and its performance, **International Journal of Man-Machine Studies**, 37, 1992, pp. 1-21.
- [Bill96] BILLINGSLEY, PATRICIA, Software Ergonomics Comes of Age: Review of HFES/ANSI Software Ergonomics Standard Content, Status, and Issues.
- [Blan91] BLANKENBERGER, S. & HAHN, K., Effects of icon design on human-computer interaction, **International Journal of Man-Machine Studies**, 35, 1991, pp. 363-377.
- [Blan00] BLANCHARD, HARRY E., Update on Human Computer Interaction Standards: Part 2: International Standards, **SIGCHI Bulletin**, July/August 2000, p. 8. (Versão *online* disponível em <http://www.acm.org/sigchi/bulletin>)
- [Bø82] BØ, KETIL, Human-Computer Interaction, **IEEE Computer**, 15(11), November 1982, pp. 9-11.
- [Boeh78] BOEHM, B. W.; BROWN, J. R.; KASPAR, H.; LIPOW, M.; MacLEOD, G. J. & MERRITT, M. J., **Characteristics of Software Quality**. North-Holland Publishing Company, Amsterdam, 1978.
- [Bond98] BOND, OLIVER, Software Usability: Quality in Use?, **Managing Information**, 5(5), June 1998. (Versão *online* disponível em www.aslib.co.uk/man-inf/jun98/article2.html).

- [Boot89] BOOTH, P., **An introduction to human computer interaction**. Lawrence Erlbaum Associates, Hillsdale, 1989.
- [Borg99] BORGHOLM, T. & MADSEN, K. H., Cooperative Usability Practices, **Communications of the ACM**, 42(5), May 1999, pp. 91-97.
- [Botm96] BOTMAN, H., Do-it-yourself usability evaluation: Guiding software developers to usability. In Jordan, P. W.; Thomas, B; Weerdmeester, B. A. and McClelland, I. L. (Eds.), **Usability evaluation in industry**. Taylor & Francis, London, 1996, pp. 59-66.
- [Brad90] BRADFORD, J. H.; MURRAY, W. D. & CARREY, T. T., What kind of errors do Unix users make? IFIP INTERACT'90 3rd International Conference on Human-Computer Interaction. 1990. **Proceedings**. Cambridge, U.K., 27-31 August, 1990, pp. 43-46.
- [Brew94] BREWSTER, S. A.; WRIGHT, P. C. & EDWARDS, A. D. N., The Design and Evaluation of an Auditory-Enhanced ScrollBar. In: ACM CHI'94 Conference, 1994. **Proceedings**. Boston - MA, 1994, pp. 173-179.
- [Brig99] BRIGNULL, HARRY, What are the advantages and disadvantages of an icon based user-interface for computer users?. **Documento Eletrônico**. www.neocortex.co.uk/essays/applied/icons.htm. (An extended essay, June 1999)
- [Broo90] BROOKE, J., BEVAN, N., BRIGHAM, E. R., HARKER, S., & YOUMANS, D., Usability assurance and standardization-work in progress in ISO. IFIP INTERACT'90 3rd International Conference on Human-Computer Interaction. 1990. **Proceedings**. Cambridge, U.K., 27-31 August, 1990, pp. 357-361.
- [Broo96] BROOKE, JOHN, SUS: a 'quick and dirty' usability scale. In Jordan, P. W.; Thomas, B; Weerdmeester, B. A. and McClelland, I. L. (Eds.), **Usability evaluation in industry**. Taylor & Francis, London, 1996, pp. 189-194.
- [Brow96] BROWN, M. G., **Keeping Score – Using the Right Metrics to Drive World-Class Performance**. Quality Resources, New York, 1996.
- [Brow98] BROWNE, DERMOT, Selling Usability Services. In: Lesley Trenner & Joanna Bawa, **The Politics of Usability**. Springer-Verlag, London - England, 1998, pp. 21-31.
- [Brun95] BRUN-COTTAN, F. & WALL, P., Using Video to Re-Present the User, **Communications of the ACM**, 38(5), May 1995, pp. 61-71.
- [Buie99] BUIE, ELIZABETH, HCI Standards: A Mixed Blessing, **Interactions**, 6(2), March-April 1999, pp. 36-42.
- [Butl85] BUTLER, K. H., Connecting theory and practice: a case study of achieving usability goals. In: ACM CHI'85 Human factors in Computer Systems, 1985. **Proceedings**. San Francisco, April 14-18, 1985, pp. 85-88.
- [Butl96a] BUTLER, K. A., Usability Engineering Turns 10, **Interactions**, 3(1), January 1996, pp. 59-75.
- [Butl96b] BUTLER, M. B. & TAHIR, M., Bringing the Users' Work to Us: Usability Roundtables at Lotus Development. In: D. R. Wixon, & J. Ramey, **Field methods Casebook for Software Design**, John Wiley & Sons – New York, 1996, pp. 249-267.
- [Buxt90] BUXTON, W., A three-state model of graphical input. In: Diaper, D., Gilmore, D., Cockton, G. & Shackel, B., **Human-Computer Interaction - INTERACT'90**, North-Holland, Amsterdam, 1990, pp. 449-456.

- [Call88] CALLAHAN, J.; HOPKINS, D.; WEISER, M. & SHNEIDERMAN, B., An empirical comparison of pie vs. linear menus. In: ACM CHI'88 Conference, 1988. **Proceedings**. Palo Alto, 1988, pp. 95-100.
- [Card78] CARD, S.; ENGLISH, W. K. & BURR, B. J., Evaluation of Mouse, Rate-controlled Isometric Joystick and Text Keys for Text Selection on a CRT, **Ergonomics**, 21, 1978, pp. 601-613.
- [Card83] CARD, S.; MORAN, T. P. & NEWELL, A., **The Psychology of Human-Computer Interaction**. Lawrence Erlbaum Associates, Hillsdale, NJ, 1983.
- [Card90] CARD, S.; MORAN, T. P. & NEWELL, A., The Keystroke-Level Model for User Performance Time with Interactive Systems. In: J. Preece, & L. Keller, **Human-Computer Interaction**, Prentice Hall International (UK) Ltd. - Hertfordshire, 1990, pp. 327-353.
- [Card94] CARD, S.; PIROLI, P. & MCKINLAY, J. D., The Cost-of-Knowledge Characteristic Function: Display Evaluation for Direct-Walk Dynamic Information Visualizations. In: ACM CHI'94 Conference, 1994. **Proceedings**. Boston - MA, 1994, pp. 238-244.
- [Carr85a] CARROLL, J. M. & ROSSON, M. B., Usability specifications as a tool in iterative development. In: H. R. Hartson (Ed.), **Advances in Human-Computer Interaction vol. 1**. Ablex, Norwood - NJ, 1985, pp. 1-28.
- [Carr85b] CARROLL, J.M. & MACK, R.L., Metaphor, computing systems and active learning, **International Journal of Man-Machine Studies**, 22, 1985, pp. 39-57.
- [Carr87] CARROLL, J. M. & ROSSON, M. B., Paradox of the Active User. In: J. M. Carroll (Ed.), **Interfacing Thought: Cognitive Aspects of Human-Computer Interaction**. The MIT Press, Cambridge, MA, 1987, pp. 80-111.
- [Carr88] CARROLL, J. M. & THOMAS, J. C., Fun, **ACM SIGCHI Bulletin**, 19(3), January 1988, pp. 21-24.
- [Carr91] CARROLL, J. M.; KELLOGG, W. A. & ROSSON, M. B., The Task-artifact Cycle. In: J. M. Carroll (Ed.), **Designing Interaction: Psychology at the Human-Computer Interface**. Cambridge University Press, Cambridge, U. K., 1991, pp. 74-102.
- [Carr95] CARROLL, J. M. & CARRITHERS, C., Managing Evaluation Goals for Training, **Communications of the ACM**, 38(7), July 1995, pp. 40-47.
- [Cast97a] CASTILLO, J.C., HARTSON, H.R. & HIX, D., Remote Usability Evaluation at a Glance. **Technical Report TR-97-12**. Virginia Tech, 1997. (Versão online disponível em miso.cs.vt.edu/~usab/remote/docs/TR_remote_evaluation.pdf)
- [Cast97b] CASTILLO, J.C., HARTSON, H.R. & HIX, D., The User-Reported Critical Incident Method at a Glance. **Technical Report TR-97-13**. Virginia Tech, 1997. (Versão online disponível em miso.cs.vt.edu/~usab/remote/docs/TR_user_reported_CI_method.pdf).
- [Cast98] CASTILLO, J.C., HARTSON, H.R. & HIX, D., Remote Usability Evaluation: Can Users Report Their Own Critical Incidents? In: ACM CHI'98 Conference, 1998. **Proceedings**. Los Angeles, CA, 18-23 April 1998, pp. 253-254.
- [Cata98] CATANI, M. B. & BIERS, D. W., Usability evaluation and prototype fidelity: users and usability professionals. In: Human Factors and Ergonomics Society 42nd Annual Meeting. **Proceedings**. 1998.

- [Cava78] CAVANO, J.P. & McCALL, J. A., A framework for the measurement of software quality. In: Software Quality and Assurance Workshop. **Proceedings**. San Diego, CA, November, 1978.
- [Chap91] CHAPANIS, A., The business case for human factors in informatics. In: Shackel, B. & Richardson, S. (Eds.), **Human Factors for Informatics Usability**. Cambridge University Press, Cambridge, UK, 1991, pp. 74-102.
- [Chig97] CHIGNELL, MARK & WATERWORTH, JOHN, Multimedia. In: Gavriel Salvendy (Ed.), **Handbook of Human Factors and Ergonomics**, pp. 1808-1861. New York: John Wiley & Sons Inc. 1997.
- [Chim94] CHIMERA, R. & SHNEIDERMAN, B., An Exploratory Evaluation of Three Interfaces for Browsing Large Hierarchical Tables of Contents, **ACM Transactions on Information Systems**, 12(4), October 1994, pp. 383-406.
- [Chin88] CHIN, J. P., DIEHL, V. A. & NORMAN, K. L., Development of an Instrument Measuring User Satisfaction of the Human-Computer Interface. In: ACM CHI'88 Conference, 1988. **Proceedings**. Washington, DC, 1991, pp. 213-218. (Versão *online* disponível em lap.umd.edu/quis/publications/chin1988.html)
- [Chri90] CHRISTIE, B. & GARDINER, M.M., Evaluation of the human-computer interface. In: J.R. Wilson & E.N. Corlett (Eds.), **Evaluation of Human Work: A Practical Ergonomics Methodology**. London: Taylor & Francis, pp. 271-320. 1990.
- [Cock93] COCKBURN, A. & GREENBERG, S., Making contact: Getting the group communicating with groupware. In: Conference on Organizational Computing Systems - COCS'93, 1993. **Proceedings**. Milpitas, CA, November 1-4 1993, pp. 31-41.
- [Cohe95a] COHEN, ANDY, A Description of MIL-STK-1472, **SIGCHI Bulletin**, 27(2), April 1995, pp. 48-49.
- [Cohe95b] COHEN, A.; CROW, D.; DILLI, I.; GORNY, P.; HOFFMAN, H.-J.; IANELLA, R.; OGAWA, K.; REITERER, H.; UENO, K. & VANDERDOCT, J., Tools for Working with Guidelines, **SIGCHI Bulletin**, 27(2), April 1995, pp. 30-32.
- [Cohi83] COHILL, A. M. & EHRICH, R. W., Automated tools for the study of human/computer interaction. In: Human Factors Society 27th Annual Meeting, 1983. **Proceedings**. Norfolk, VA, October 10-14 1983, pp. 897-900.
- [Cons92] CONSTANTINE, L. L., Going to the Source, **Computer Language**, December 1992, pp. 110-112.
- [Cons94] CONSTANTINE, L. L., Collaborative Usability Inspections for Software. In: Software Development '94. **Proceedings**. London, U.K., 4-7 September 1984, pp. 269-274.
- [Cons99] CONSTANTINE, LARRY L. & LOCKWOOD, LUCY A. D., **Software for Use: A Practical Guide to the Models and Methods of User-Centered Design**, ACM Press (New York - NY)/Addison-Wesley Longman Inc. (Reading - MA), 1999.
- [Cost99] COSTABILE, M. F. & MATERA, M., Evaluating WIMP Interfaces through the SUE Approach. In: 10th International Conference on Image Analysis and Processing, ICIAP'98. **Proceedings**. Venice, Italy, 27 - 29 September, 1999, pp. 1192-1197.
- [Cout88] COUTAZ, J., De l'ergonome à l'informaticien: pour une méthode de conception et de réalisation des systèmes interactifs. Colloque ERGO-IA-AFCET-SELF. **Actes**. Biarritz, France, 4-6 Octobre, 1988.

- [Cox93] COX, K. & WALKER, D., **User-interface Design**, Prentice Hall/Simon & Schuster (Asia) Pte. Ltd. - Singapore, 1993.
- [Cox94] COX, M. E.; O'NEAL, P. & PENDLEY, W. L., UPAR Analysis: Dollar Measurement of a Usability Indicator for Software Products. In: R. G. Bias & D. J. Mayhew (Eds.), **Cost-Justifying Usability**. Academic Press, Chesnut Hill, Massachusetts, 1994, pp. 145-158.
- [Crer98] CRERAR, ALISON & BENYON, DAVID, Integrating Usability into Systems Development. In: Lesley Trenner & Joanna Bawa, **The Politics of Usability**. Springer-Verlag, London - England, 1998, pp. 49-60.
- [Cres94] CRESWELL, J., **Research design: Qualitative and quantitative approaches**. Sage, London, 1994.
- [Croc93] CROCKER, D., Making Standards the IETF Way, **StandardView**, 1(1), September 1993, p. 48.
- [Cros92] CROSBY, P. B., **Quality Is Free: The Art of Making Quality Certain**. Mentor Books, New York, 1992.
- [Crow95] CROW, DANIEL, HalClon - Designing for People, **SIGCHI Bulletin**, 27(2), April 1995, pp. 40-41.
- [Cuom92] CUOMO, D. L. & BOWEN, C. D., Stages of user activity model as a basis for user-system interface evaluations. In: Human Factors Society 36th Annual Meeting, 1992. **Proceedings**. Atlanta, GA, October 12-16 1992, pp. 1254-1258.
- [Curt93] CURTIS, BILL, Maturity From the Users Point of View, **IEEE Software**, 26(7), July 1993, pp. 89-90.
- [Curt94] CURTIS, B. & HEFLEY, B., A WIMP no more, **Interactions**, 1(1), January 1994, pp. 22-34.
- [Cyert63] CYERT, R. & MARCH, J., **A Behavioral Theory of the Firm**, Prentice Hall, Inc., Englewood Cliffs, NJ, 1963.
- [Davi93] DAVIES, C. & MEDYCKYJ-SCOTT, D., The USIS Project: Surveying User Opinion on GIS User Interfaces. In: Fourth European Conference on Geographical Information Systems, 1993. **Proceedings**. Genoa, Amsterdam, 29 March - 1 April, 1993, pp. 476-483.
- [Dele92] DELEN, G. P.A. J. & RIJSENBRIJ, D. B. B., The Specification, Engineering and Measurement of Information Systems Quality, **The Journal of Systems and Software**, 17(1992), pp. 205-217.
- [Dees79] DEESE, D., Experiences measure user satisfaction. In: ACM Computer Measurement Group 1979. **Proceedings**. Dallas, December 1979.
- [delG96] del GALDO, E. & NIELSEN, J. (Eds.), **International User Interfaces**. John Wiley & Sons, New York, NY, 1996.
- [Delo92] DELONE, W. H. & MCLEAN, E. R. Information systems success: The quest for the dependent variable, **Information Systems Research**, 3(1), March 1992, pp. 60-95.
- [Demi86] DEMING, W. EDWARDS, **Out of the Crisis**. MIT Center for Advanced Engineering Study, Cambridge, MA, 1986.
- [Denn90] DENNING, S.; HOIEM, D.; SIMPSON, M. & SULLIVAN, K., The value of thinking-aloud Protocols in industry: A case study at Microsoft Corporation. In: Human Factors Society 34th Annual Meeting, 1990. **Proceedings**, pp. 1285-1289.

- [Denn92] DENNING, PETER J., **What is Software Quality?** A Commentary from Communications of ACM, January 1992. **Documento Eletrônico**. *cnø.gmu.edu/pjd/PUBS/softqual92.pdf*.
- [Desu91] DESURVIRE, H.; LAWRENCE, D. & ATWOOD, M., Empiricism versus judgement: Comparing user interface evaluation methods on a new telephone-based interface, **SIGCHI Bulletin**, 23(4), October 1991, pp. 58-59.
- [Desu92] DESURVIRE, H.; KONDZIELA, J. & ATWOOD, M., What Is Gained And Lost When Using Evaluation Methods Other Than Empirical Testing. In: ACM CHI'92 Conference, 1992. **Proceedings**. Monterey, California, May 3-7 1992, collection of abstracts, pp. 125-126.
- [Desu94] DESURVIRE, H. W., Faster, Cheaper!! Are Usability Inspection Methods as Effective as Empirical Testing?. In: Jakob Nielsen and Robert L. Mack, **Usability Inspection Methods**. John Wiley & Sons, Inc. New York, 1994, pp. 173-202.
- [Deut87] DEUTSCH, M. S. & WILLIS, R. R., **Software Quality Engineering**. Prentice Hall, 1987.
- [Diel94] DIELI, M.; DYE, K.; MCCLINTOCK, M. & SIMPSON, M., The Microsoft Corporation Usability Group. In: M. E. Wiklund (Ed.) **Usability in Practice**. Academic Press, Toronto, 1994, pp. 327-357.
- [Dill95] DILLI, I. & HOFFMAN, H.-J., HyperSAM - A Management Tool for Large User interface Guideline Sets, **SIGCHI Bulletin**, 27(2), April 1995, pp. 42-45.
- [DIN88] DEUTSCH INSTITUT FÜR NORMUNG - DIN, **DIN 2137, German keyboard for typewriters - Allocation of Characters to Keys** (Parts 1, 2, 6 and 11). 1988.
- [Dix98] DIX, A., FINLAY, J., ABOARD, G. & BEALE, R., **Human-Computer Interaction**. Prentice Hall Europe, Hertfordshire, 1998, 2nd edition.
- [Dix99] DIX, A., Design of User Interfaces for the Web. In: User Interfaces to Data Intensive Systems UIDIS'99. Edinburgh, Scotland, 5 - 6 September 1999, pp. 2-11.
- [Dixo96] DIXON M. B., COXHEAD J. F. & DODMAN E. A., The Enhancement of CASE to Improve Software Quality. **Documento eletrônico**. *www.lmu.ac.uk/ies/comp/staff/mdixon/Research/icsq.html*.
- [Doll76] DOLOTTA, T. A.; BERNSTEIN, R. S.; DICKSON, S. (Jr); FRANCE, N. A.; ROSENBLATT, B. A.; SMITH, D. M. & STEEL, T. B. (Jr), **Data Processing in 1980-1985, A Study of Potential Limitations to Progress**. Wiley-Interscience, NY. 1976.
- [Doll88] DOLL, W.J. & TORKZADEH, G., The Measurement of End-User Computing Satisfaction, **MIS Quarterly**, 12(2), June 1988, pp. 259-276.
- [Dori93] DORLING, A., SPICE: Software Process Improvement and Capability Determination, **Software Quality Journal**, 2 (1993), pp. 209-224.
- [Doub97] DOUBLEDAY, A.; RYAN, M.; SPRINGETT, M. & SUTCLIFFE, A., A comparison of usability techniques for evaluating design. In: Conference on Designing Interactive systems: processes, practices, methods, and techniques. **Proceedings**. The Netherlands, August 18 - 20, 1997, pp. 101-110.

- [Doug99] DOUGLAS, S. A.; KIRKPATRICK, A. E. & MACKENZIE, I. S., Testing pointing device performance and user assessment with the ISO9241, Part 9 standard. In: CHI'99 Conference on Human Factors in Computing Systems: the CHI is the limit. **Proceedings**. Pittsburgh, PA, May 15 - 20 1999, pp. 215-222.
- [Down91] DOWNTON, ANDY, Evaluation Techniques for Human-Computer Systems Design. In: Andy Downton (Ed.), **Engineering the Human-Computer Interface**. McGraw-Hill Book Company (UK) Ltd. - London, 1991, pp. 325-355.
- [Dray00] DRAY, SUSAN, Thoughts from 35,000 feet: The evolving real-world context of user centered design, **Interactions**, 7(3), May/June 2000, pp. 22-26.
- [Drea98] DREACHELIN, J. L., Conducting Effective Focus Groups in the Context of Diversity: Theoretical Underpinnings and Practical Implications, **Qualitative Health Research**, 8(6), November 8 1998, pp. 813-821.
- [Drom96] DROMEY, R. G., Cornering the Chimera, **IEEE Software**, 29(1), January 1996, pp. 33-43.
- [Duma89] DUMAS, J. S., Stimulating Change Through Usability Testing, **ACM SIGCHI Bulletin**, 21 (1), July 1989, pp. 37-44.
- [Duma94] DUMAS, J. S. & REDISH, J. C., **A Practical Guide to Usability Testing**. Ablex Publishing Co., Norwood, NJ, second printing 1994.
- [Duma95] DUMAIS, SUSAN, Finding What You Want: New Tools and Tricks, **IEEE Software**, 28(10), October 1995, pp. 79-81 e 86.
- [Dzid78] DZIDA, W.; HERDA, S. & ITZFELDT, W. D., User-Perceived Quality of Interactive Systems, **IEEE Transactions on Software Engineering**, SE-4(4), July 1978, pp. 270-276.
- [Dzid89] DZIDA, WOLFGANG, The Development of Ergonomic Standards, **SIGCHI Bulletin**, 20(3), January 1989, pp. 35-43.
- [Dzid95] DZIDA, WOLFGANG, Standards for user-interfaces, **Computer Standards Interfaces**, 17, January 1995, pp. 89-97.
- [Dzid96] DZIDA, WOLFGANG, The Development of Ergonomic Standards, **ACM Computing Surveys**, 28(1), March 1996, pp. 173-175.
- [Easo84] EASON, K. D., Towards the experimental study of usability, **Behaviour & Information Technology**, 3(2), March/April 1984, pp. 133-143.
- [EB99] ENCYCLOPÆDIA BRITANNICA, Engineering. **Documento Eletrônico**. www.britannica.com/bcom/eb/article/7/0,5716,108127+1+105842,00.html.
- [Eber93] EBERTS, R. E. & BITTIANDA, K. P., Preferred mental models for direct manipulation and command-based interfaces, **International Journal of Man-Studies**, 38, pp. 769-785.
- [Eber97] EBERTS, RAY, Cognitive Modelling. In: Gavriel Salvendy (Ed.), **Handbook of Human Factors and Ergonomics**, pp. 1328-1374. New York: John Wiley & Sons Inc. 1997.
- [Ehr194a] EHRLICH, KATE & ROHN, J. A., Cost Justification of Usability Engineering: A Vendor's Perspective. In: R. G. Bias & D. J. Mayhew (Eds.), **Cost-Justifying Usability**. Academic Press, Chesnut Hill, Massachusetts, 1994, pp. 73-110.
- [Ehr194b] EHRLICH, K.; BUTLER, M. B. & PERNICE, K., Getting the Whole Team into Usability Testing, **IEEE Software**, 27(1), January 1994, pp. 89-91.

- [EL00] ERGOLIGHT, Ergolight Lab Tester. Documento Eletrônico. www.ergolight.co.il/www/LabTester.
- [Elgi95] ELGIN, B., Subjective Usability Feedback from the Field over the Network, **SIGCHI Bulletin**, 27(4), October 1995, pp. 43-44. (Versão online disponível em www.acm.org/sigchi/chi95/proceedings/sigs/be_bdy.htm.)
- [Enge99] ENGEL, FRED, Developing Standards for IT, **International Journal of Network Management**, 9(5), 1999, pp. 303 – 304.
- [EOQ00] EUROPEAN ORGANISATION FOR QUALITY – EOQ, Towards a European Vision of Quality – The way forward. **Report version 1.0**, Helsinki, January 2000.
- [Epps86] EPPS, B. W., Comparison of Cursor Control Devices Based on Fitt's Law Models. In: Human Factors Society 30th Annual Meeting. **Proceedings**. Dayton, OH , 1986, pp. 327-331.
- [Ereb94] EREBACK, ANNA-LENA & HÖÖK, KRISTINA, Using Cognitive Walkthrough for Evaluating a CSCW Application. In: ACM CHI'94 Conference, 1994. **Proceedings**. Boston - MA, 1994, pp. 91-92.
- [Eric84] ERICSSON, K. A. & SIMON, H. A., **Protocol Analysis: Verbal Reports as Data**. The MIT Press, Cambridge, MA, 1984.
- [Este90] ESTEVAM, R. DE C., **O Estudo Sobre Desenvolvimento de Interfaces: Definição de Técnicas de Classificação e de Avaliação Baseadas na Satisfação do Usuário**. (Dissertação de Mestrado) - COPPE/UFRJ, Rio de Janeiro, Abril, 1990.
- [ETSI93] European Telecommunications Standards Institute - ETSI, **ETSI ETR 095 Human Factors: Guide for Usability Evaluations of Telecommunications Systems and Services**. Sophia Antipolis, France. 1993.
- [Evan76] EVANS, J., Measures of Computer and Information Systems Productivity Key Informant Interviews. **Technical Report APR-20546/TR-5**. Westinghouse Research Laboratories, Pittsburgh, PA, October 1976.
- [Fair85] FAIRLEY, R. **Software Engineering Concepts**. McGraw-Hill, New York, 1985.
- [Fari96] FARIAS, GIOVANNI F. DE, **Diretrizes para Projeto de Interfaces Homem-Máquina Aplicadas a Sistemas de Supervisão de processos Industriais**. (Dissertação de Mestrado em Informática) - DCC/CCT/UFPB, Campina Grande, Abril, 1996.
- [Fede99] FEDERICO, MARCELLO, Usability Evaluation of a Spoken Data-Entry Interface. In: IEEE International Conference on Multimedia Computing and Systems ICMCS'99. Florence, Italy. 7 - 11 June 1999, Volume I, pp. 726-731.
- [Feig91] FEIGENBAUM, A. V., **Total Quality Control**, McGraw-Hill, New York, 1991, 3rd edition (40th anniversary edition).
- [Feld97] FELDT, TARU, The role of sense of coherence in well-being: the analysis of main and moderator effects, **Work & Stress**, 11(2), August 1997, pp 134-147.
- [Fent96] FENTON, N. E. & PFLEEGER, S. L., **Software Metrics: A rigorous and practical approach**, Thomson Computer Press, 1996.
- [Fern95] FERNANDEZ, D. R. & PERREWÉ, P. L., Implicit Stress Theory: an experimental examination of subjective performance information on employee evaluations, **Journal of Organizational Behavior**, 16, pp. 353-362.

- [Flow90] FLOWER, JOE, Managing Quality: A discussion with David Garvin, **Healthcare Forum Journal**, 33(5), September-October 1990. (Versão *online* disponível em www.well.com/user/bbear/garvin.html)
- [Fole82] FOLEY, J. D., & VAN DAM, A., **Fundamentals of Interactive Computer Graphics**. Addison-Wesley, Reading, MA, 1982.
- [Fole84] FOLEY, J. D.; WALLACE, V. L. & CHAN, P., The Human Factors of Computer Graphics Interaction Technique, **IEEE Computer Graphics and Applications**, 4(11), November 1984, pp. 13-48.
- [Ford90] FORD, GARY, SEI Report on Undergraduate Software Engineering Education. **Technical Report CMU/SEI-90-TR-003 ESD-TR-90-204**. March, 1990.
- [Fran91] FRANZKE, M., Evaluation Technique Evaluated: Experience Using the Cognitive Walkthrough. Bellcore/BCC Symposium on User-Centered Design. **Proceedings**. Livingston, NJ, November 1991.
- [Fran93] FRANZKE, M.; MARX, A. N.; ROBERTS, T. L. & ENGELBECK, G. E., Is Speech Recognition Usable? An Exploratory Study of the Usability of a Speech-Based Voice Mail Interface, **SIGCHI Bulletin**, 25(3), July 1993, pp. 49-51.
- [Fran95] FRANZKE, M., Turning Research into Practice: Characteristics of Display-Based Interaction. In: ACM CHI'95 Conference on Human Factors in Computing Systems. **Proceedings**. Denver, Colorado, May 7-11 1995, pp. 421-428.
- [Frat99] FRATERNALI, PIERO, Tools and approaches for developing data-intensive Web applications: a survey, **ACM Computing Surveys**, 31(3), September 1999, pp. 227-263.
- [Fred99] FREDERICKSON-MELE, K & CONRAD, F. Report on the ARIES 2 Cognitive Walkthrough. **Documento Eletrônico**. cpcug.org/user/hamilev/chi99/Mele%20and%20Conrad.htm.
- [Fres91] FRESE, M.; BRODBECK, F. C.; ZAPF, D. & PROMPER, J., Users Errors and Error Handling: Its Relationships with Task Structure and Social Support, **SIGCHI Bulletin**, 23(2), April 1991, pp. 59-62.
- [Fröl00] FRÖHLICH, BERND & PLATE, JOHN, The Cubic Mouse: A New Device for Three-Dimensional Input. In: CHI 2000 Conference on Human Factors in Computing Systems. **Proceedings**. The Hague, Netherlands, April 1 - 6 2000, pp. 526-531.
- [Fuhr98] FUHRMANN, ANTON & GRÖLLER, EDUARD, Real-time techniques for 3D flow visualization. In: Conference on Visualization '98. **Proceedings**. Research Triangle Pk, USA, October 18 - 23 1998, pp. 305-312.
- [Gabb99] GABBARD J. L.; HIX, D. & SWAN II, J. E., User-Centered Design and Evaluation of Virtual Environments, **IEEE Computer Graphics and Applications**, 19(6), November/December 1999, pp. 51-59.
- [Gabl94] GABLE, G., Integrating Case Study and Survey Research Methods: An Example in Information Systems, **European Journal of Information Systems**, 3(2), 1994, pp. 112-126.
- [Gain83] GAINES, B. R. & SHAW, M. L. G., Dialog Engineering. In: M. E. Sime & M. J. Coombs (Eds.), **Designing for Human-Computer Communication**. London, 1983, pp. 23-53.
- [Gale87] GALE, A. & CHRISTIE, B. (Eds.), **Psychophysiology and the Electronic Workplace**. John Wiley and Sons Ltd, Chichester, 1987.

- [Gale89] GALE, A., Adding audio and video to an office environment. In: First European Conference on Computer Supported Co-operative Work, 1989. **Proceedings**. The Hilton Hotel, Gatwick, London, UK, 13-15 September 1989, pp. 121-130.
- [Gale96] GALE, STEPHEN, A collaborative approach to developing style guides. In: Conference on Human Factors in Computing Systems. **Proceedings**. Vancouver, Canada, April 13 - 18 1996, pp. 362-367.
- [Gali85] GALITZ, W. O., **Handbook of Screen Design**. QED Information Sciences. North-Holland, 1985.
- [Gall91] GALLUPE, R. B., BASTIANIUTTI, L. M. & COOPER, W. H., Unblocking brainstorming, **Journal of Applied Psychology**, 76 (1), January 1991, pp. 137-142.
- [Gard86] GARDINER, M. M., Psychological issues in adaptive interface design. In: IEE Colloquium on Adaptive Interface Design Organised by Professional Group C5 (Man-Machine Interaction). **Proceedings**. Savoy Place, London, 4 November 1986. IEE Digest No. 1986/110, 6/1-6/3.
- [Garv84] GARVIN, D. A., What Does "Product Quality" Really Mean?, **Sloan Management Review**, Fall 1984, pp. 25-45.
- [Garv88] GARVIN, D. A., **Managing Quality**. Free Press, New York, 1988.
- [Garv94] GARVIN Quality Framework. **Documento Eletrônico**. <http://www.tantara.ab.ca/qualframe.htm>
- [Gent96] GENTLEMAN, W. M., If software quality is a perception, how do we measure it? In: Sixth International Conference on Software Quality (6ICSQ), 1996. **Proceedings**. Ottawa, Canada, 28-30 October, 1996, pp. 335-345. (Versão *online* disponível em wwwsel.it.nrc.ca/seldocs/cpdocs/NRC40149.pdf)
- [Gedi99] The IsoMetrics usability inventory: An operationalisation of ISO 9241/10, **Behavior and Information Technology**, 18, 1999, pp. 151-164. (Documento eletrônico disponível em <http://www.infj.ulst.ac.uk/~ccc23/papers/isometer.html>)
- [Gerh96] GERHARDT-POWALS, J., Cognitive engineering principles for enhancing human-computer performance, **International Journal of Human-Computer Interaction**, 8(2), 1996, pp. 189-211.
- [Gibb94] GIBBS, W. W., Software's Chronic Crisis, **Scientific American**, September 1994, pp. 72-81.
- [Gilb83] GILBERT, P., **Software Engineering**. Winthrop, Cambridge, 1983.
- [Gilb84] GILB, T., The "impact analysis table" applied to human factors design. In: Interact'84 1st IFIP Conference on Human-Computer Interaction, 1984. **Proceedings**. London - UK, September 4-7, 1984, vol. 2, pp. 97-101.
- [Gilb87] GILB, T., **Principals of Software Engineering Management**. Addison-Wesley, Reading, MA, 1987.
- [Gill93] GILLAN, D. J., A Componential Model of Human Interaction With Graphical Displays, **SIGCHI Bulletin**, 25(3), July 1993, pp. 64-66.
- [Gill92] GILLIES, ALAN C., **Software Quality, Theory and Management**. Chapman & Hall, 1992.
- [Giss98] GISSELQUIST, RICHARD, Engineering in Software, **Communications of the ACM**, 41(10), October 1998, pp. 107-108.

- [Glas97] GLASS, ROBERT L., Telling Good Numbers from Bad Ones, **IEEE Software**, July/August 1997, pp. 15-16 e 19.
- [Gold93] GOLDBERG, D., RICHARDSON, G., Touch-typing with a stylus. In: INTERCHI'93 - CHI'93 & INTERACT'93 Conference. **Adjunct Proceedings**. Amsterdam, The Netherlands, 1993, pp. 80-87.
- [Gold99] GOLDSTEIN, M.; BOOK, R.; ALSIÓ, G. & TESSA, S., Non-keyboard QWERTY touch typing: a portable input interface for the mobile user. CHI'99 Conference on Human Factors in Computing Systems: the CHI is the limit. **Proceedings**. Pittsburgh, PA, May 15 - 20 1999, pp. 32-39.
- [Gome99] GOMES, PERICLES V., Usability Feedback in Educational Interfaces, **Educational Technology & Society**, 2(4), 1999. (Versão online disponível em ifets.gmd.de/periodical/vol_4_99/pericles_gomes_website_review.html).
- [Gong94] GONG, R. & KIERAS, D., A Validation of the GOMS Model Methodology in the Development of a Specialized, Commercial Software Application. In: ACM CHI'94 Conference, 1994. **Proceedings**. Boston - MA, 1994, pp. 351-357.
- [Good85] GOOD, MICHAEL, The use of logging data in the design of a new text editor. In: ACM CHI'85 Conference, 1985. **Proceedings**. New York, 1985, pp. 1032-1043.
- [Good86] GOOD, M.; SPINE, T. M.; WHITSIDE, J. & GEORGE, P., User-derived impact analysis as a tool for usability engineering. ACM CHI'86 Conference, 1986. **Proceedings**. ACM Press, New York, 1986, pp. 241-246.
- [Good89] GOOD, MICHAEL, Developing the XUI style. In: Jacob Nielsen (Ed.), **Coordinating User Interfaces for Consistency**. Academic Press, Boston, MA, 1989, pp. 57-73.
- [Good87] GOODWIN, N. C., Functionality and Usability, **Communications of the ACM**, 30(3), March 1987, pp. 229-233.
- [Gorn95] GORNY, PETER, EXPOSE - An HCI-Counseling Tool for User Interface Designers, **SIGCHI Bulletin**, 27(2), April 1995, pp. 35-37.
- [Goul85] GOULD, J. D. & LEWIS, C., Designing for Usability: Key Principles and What Designers Think, **Communications of the ACM**, 28(3), March 1985, pp. 300-311.
- [Gray93] GRAY, J. S., Menu Stacking - Help or Hindrance?, **SIGCHI Bulletin**, 25(3), July 1993, pp. 55-57.
- [Gree91] GREENBAUM, J. & KYNG, M. (Eds.), **Cooperative Design of Computer Systems**. LEA, Hillsdale - NJ, 1991.
- [Gree92] GREENBERG, S.; ROSEMAN, M.; WEBSTER, D. & BOHNET, R., Human and technical factors of distributed group drawing tools, **Interacting with Computers**, 4(3), March 1992, pp. 364-392.
- [Grud89] GRUDIN, J., The Case Against User Consistency, **Communications of the ACM**, 32(10), October 1989, pp. 1164-1173.
- [Grud90] GRUDIN, J., The Computer Reaches Out: The Historical Continuity of Interface Design. In: CHI'90 Conference. **Proceedings**. Seattle, 1990, pp. 261-268.
- [Grud91] GRUDIN, J., Interactive Systems: Bridging the Gaps Between Developers and Users, **IEEE Computer**, 24(4), April 1991, pp. 59-69.

- [Grud92] GRUDIN, J., Utility and usability: research issues and development contexts, **Interacting with Computers**, 4(2), 1992, pp. 209-217.
- [Grud94] GRUDIN, J., Groupware and social dynamics: Eight challenges for developers, **Communications of the ACM**, 37(1), January 1994, pp. 92-105.
- [Grud95] GRUDIN, J. & PALEN, L., Why Groupware Succeeds: Discretion or Mandate? In: ECSCW'95 Conference. Proceedings. Seattle, 1990, pp. 261-268.
- [Gunn98] GUNN, C., System Log Data. In: Jen Harvey (Ed.), **Evaluation Cookbook**. LTDI - Institute for Computer Based Learning/Heriot-Watt University, Edinburgh, 1998, p. 62.
- [Guti00] GUTIERREZ, P. & RITZIE, S., Assessing the User Experience: The Role of Usability in Internet Services. **Documento Eletrônico**. www.luminant.com/IMAGES/WP_usability.pdf.
- [Habe91] HABERMANN, FRITZ, Giving Real Meaning to 'Easy-to-use' Interfaces, **IEEE Software**, 24(7), July 1991, pp. 90-91.
- [Hack92] HACKMAN, G. S. & BIERS, D. W., Team usability testing: Are two heads better than one?. In: Human Factors Society Annual Meeting, 1992. **Proceedings**. 1992, pp. 1205-1209.
- [Hack98] HACKOS, J. T. & REDISH, J. C., **User and Task Analysis for Interface Design**. John Wiley & Sons, Inc. New York, N.Y., 1998.
- [Hail91] HAILSTONE, R., Quality Management and Software Engineering. In: Darrel Ince (Ed.): **Software Quality and Reliability, Tools and methods**, Chapman & Hall, London. 1991.
- [Hamm94] HAMMONTREE, M.; WEILER, P. & NAYAK, N., Remote Usability Testing, **Interactions**, 1(7), July 1994, pp. 21-25.
- [Hans97] HANSELMAN, DUANE & LITTLEFIELD, BRUCE, **The Student Edition of MATLAB – Version 5 User's Guide**, Prentice Hall, The MATLAB Curriculum Series, Upper Saddle River – NJ, 1997.
- [Hans84] HANSON, G. C.; KRAUT, R. E. & FARBER, J. M., Interface design and multivariate analysis of UNIX command use, **ACM Transactions on Office Information Systems**, 2, 1984, pp. 42-57.
- [Harp93] HARPER, B. D. & NORMAN, K. L., Improving User Satisfaction: The Questionnaire for User Interaction Satisfaction Version 5.5. In: 1st Annual Mid-Atlantic Human Factors Conference. **Proceedings**. Virginia Beach, VA, 1993, pp. 224-228. (Versão *online* disponível em lap.umd.edu/quis/publications/harper1993.html)
- [Harp97] HARPER, B. D.; SLAUGHTER, L. A. & NORMAN, K. L., Questionnaire Administration via the WWW: A Validation & Reliability Study for a User Satisfaction Questionnaire. In: WebNet97 (Association for the Advancement of Computing in Education). **Proceedings**. Toronto, Canada, November 1997. (Versão *online* disponível em <http://www.lap.umd.edu/webnet/paper.html>)
- [Harr98] HARRISON, F.; GUJAR, A.; MOCHON, C. & WANT, R., Squeeze Me, Hold Me, Tilt Me! An Exploration of Manipulative User Interfaces. In: CHI'98 Conference. **Proceedings**. Los Angeles, CA, 18-23 April 1998, pp. 17-24.
- [Hart90] HARTSON, H. R.; SIOCHI, A. C. & HIX, D., The UAN: A User-Oriented Representation for Direct Manipulation Interface Designs, **ACM Transactions on Information Systems**, 8(3), July 1990, pp. 181-203.

- [Hart96] HARTSON, H. R., CASTILLO, J. C., KELSO, J., KAMLER, J. & NEALE, W. C., Remote Evaluation: The Network as an Extension of the Usability Laboratory. In: ACM CHI'96 Conference, 1996. **Proceedings**. Monterey, 1996, pp. 251-258.
- [Hart98] HARTSON, H. R. & CASTILLO, J. C., Remote evaluation for post-deployment usability improvement. In: Working Conference of Advanced Visual Interfaces AVI'98. **Proceedings**. 1998, pp. 22-29.
- [Harv98] HARVEY, JEN (Ed.), **Evaluation Cookbook**. LTDI - Institute for Computer Based Learning/Heriot-Watt University, Edinburgh, 1998.
- [Hayh90] HAYHOE, D., Sorting-based menu categories, **International Journal of Man-Machine Studies**, 33, 1990, pp. 677-695.
- [HCI96] HCI, Special Issue on Current Perspectives on Participatory Design, **Human-Computer Interaction**, Vol. 11, 1996, pp. 181-290.
- [Hell96] HELLER, HAGAN & RUBERG, ALAN, Usability Studies on a Tight Budget - Guerrilla Labs. **Documento eletrônico**. <http://www-pcd.stanford.edu/asd/info/articles/guerrilla-labs.html>. 1996.
- [Henn95] HENNINGER, S.; HAYNES, K. & REITH, M, W., A framework for developing experience-based usability guidelines. Conference on Designing Interactive Systems: Processes, Practices, Methods, & Techniques. **Proceedings**. Ann Arbor, MI, August 23 - 25 1995, pp. 43-53.
- [Henr98] HENRY, P., **User-Centered Information Design for Improved Software Usability**. Artech House, Inc., Norwood - MA, 1998.
- [Hewe87] HEWETT, T. T. & SCOTT, S., The use of thinking-out-loud and protocol analysis in development of a process model of interactive database searching. In: IFIP INTERACT'87 Second International Conference on Human-Computer Interaction, 1987. **Proceedings**. Stuttgart, Germany, 1-4 September 1987, pp. 51-56.
- [Hibi97] HIBINO, S. & RUNDENSTEINER, E. A., User Interface Evaluation of a Direct Manipulation Temporal Visual Query Language. In: 5th ACM International Conference on Multimedia. **Proceedings**. Seattle, Washington, November 9 - 13 1997, pp. 99-107.
- [Hill97] HILL, W. C. & TERVEEN, L. G., Involving remote users in continuous design of web content. In: Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques. **Proceedings**. The Netherlands, August 18 - 20, 1997, pp. 137-145.
- [Hilt81] HILTZ, S. R. & TUROFF, M., The evolution of user behavior in a computer conference system, **Communications of the ACM**, 24(11), November 1981, pp. 739-751.
- [Hinc99a] HINCLEY, KEN & SINCLAIR, MIKE, Touch-Sensing Input Devices. CHI'99 Conference on Human Factors in Computing Systems: the CHI is the limit. **Proceedings**. Pittsburgh, PA, May 15 - 20 1999, pp. 223-230.
- [Hinc99b] HINCLEY, K.; SINCLAIR, M.; HANSON, E.; SZELISKI, R. & CONWAY, M., The VideoMouse: A Camera-Based Multi-Degree-of-Freedom Input Device. In: 12th Annual ACM Symposium on User Interface Software and Technology UIST'99. **Proceedings**. Asheville, NC, November 7 - 10, 1999, pp. 103-112.
- [Hirs95] HIRSCHMAN, L. & CUOMO, D., Evaluation of Human Computer Interfaces: A Report from an ARPA Workshop, **SIGCHI Bulletin**, 27(2), April 1995, pp. 28-29.

- [Hix91] HIX, D. & SCHULMAN, R. S., Human-Computer Interface Development Tools: A Methodology for Their Evaluation, **Communications of the ACM**, 34(3), March 1991, pp. 74-87.
- [Hix93] HIX, D. & HARTSON, H. R., **Developing User Interfaces: Ensuring Usability Through Product & Process**. John Wiley & Sons Inc., New York, 1993.
- [Hix99] HIX, D.; SWAN II, J. E.; GABBARD, J. L.; MCGEE, M.; DURBIN, J. & KING, T., User-Centered Design and Evaluation of a Real-Time Battlefield Visualization Virtual Environment. In: IEEE Virtual Reality VE'99. **Proceedings**. Houston, Texas, 13 - 17 March 1999, pp. 96-103.
- [Hold89] HOLDAWAY, K. & BEVAN, N., User System Interaction Standards, **Computer Communications**, 12(2), April 1989, pp. 92-102.
- [Holl86] HOLLANDA, A. B. DE, **Novo Dicionário Aurélio da Língua Portuguesa**. Editora Nova Fronteira, São Paulo, Brasil., 2ª edição (revista e ampliada), 1986.
- [Horn90] HORN, A. S., **Oxford Advanced Learner's Dictionary**. Oxford Press University, Oxford, UK, 4th edition, 1990.
- [Horn97] HORNOF, ANTHONY J. & KIERAS, DAVID E., Cognitive modeling reveals menu search in both random and systematic. In: Conference on Human Factors in Computing Systems. **Proceedings**. Atlanta, GA, March 22 - 27 1997, pp. 107-114.
- [Horn99] HORNOF, ANTHONY J. & KIERAS, DAVID E., Cognitive modeling demonstrates how people use anticipated location knowledge of menu items. In: CHI'99 Conference on Human Factors in Computing Systems: the CHI is the limit. **Proceedings**. Pittsburgh, PA, May 15-20 1999, pp. 410-417.
- [Hump89] HUMPHREY, W. S., **Managing the software process**, Addison-Wesley, 1989.
- [Hutc85] HUTCHINS, E.L.; HOLLAN, J.D. & NORMAN, D.A., Direct manipulation interfaces. **Human-Computer Interaction**, 1(4), 1985, pp. 311-338
- [Hyat96] HYATT, L. & ROSENBERG, L. A Software Quality Model and Metrics for Identifying Project Risks and Assessing Software Quality. **Documento Eletrônico**. <http://www.tantara.ab.ca/qualmodel.htm>. 1996.
- [Iann95] IANNELLA, RENATO, Diades II - A Multi-Agent Interface Design Approach with an Integrated Assessment Component, **SIGCHI Bulletin**, 27(2), April 1995, pp. 33-34.
- [IBM91a] IBM INC., **Guide to User Interface Design**, Nr. SC34-4289, IBM Inc., 1991.
- [IBM91b] IBM INC., **Advanced Interface Design Guide**, Nr. SC26-4582, IBM Inc., 1991.
- [IBM99] IBM, The ANSI/HFES 200 standard. **Documento Eletrônico**. <http://www.ibm.com>. 1999.
- [IEC00] INTERNATIONAL ELECTROTECHNICAL COMMISSION, Inside the IEC - General Information. **Documento Eletrônico**. www.iec.ch/gnote1-e.htm.
- [IEEE 90a] INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, **IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries**. New York - NY, 1990. (Versão online disponível em www.sei.cmu.edu/str/indexes/glossary/usability.html).
- [IEEE 90b] INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, **IEEE Std 610.12 - Standard Glossary of Software Engineering Terminology (ANSI)**. New York, 1990.

- [IEEE92] INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, **IEEE Std P1061 – Standard for a Software Quality Metrics Methodology**. New York, 1992.
- [Ilg91] ILG, R.; KRONEBERG, M & RICHTER, K., **Checkliste für die Evaluation von Leitständen**. Fraunhofer-Institut für Arbeitswirtschaft und Organisation, Stuttgart, 1991.
- [INUS97] INUSE, **Usability Standards. Documento eletrônico**. http://www.lut.ac.uk/research/husat/inuse/f_usability_standards.html. May, 1997 (última atualização).
- [ISC97] INFRASTRUCTURE STEERING COMMITTEE, COMPUTER SCIENCE AND TELECOMMUNICATIONS BOARD, COMMISSION ON PHYSICAL SCIENCES, MATHEMATICS, AND APPLICATIONS, **More than Screen Deep: Toward Every-Citizen Interfaces to the Nation's Information Infrastructure Information**. National Research Council National Academy Press, Washington, D.C., 1997.
- [ISO86] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, **ISO/IEC9126 – Quality – Vocabulary**. Geneva, 1986.
- [ISO87] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, **ISO 9000 – Quality Management and Quality Assurance Standards – Guidelines for selection and use**. Geneva, 1987.
- [ISO92] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, **ISO 9126 – Information Technology - Software Product Evaluation - Quality Characteristics and Guidelines for Their Use**. Geneva, 1992.
- [ISO97a] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, **ISO 9241 Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 14: Menu Dialogues**. Geneva, 1997.
- [ISO97b] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, **ISO 9241 Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 15: Command Dialogues**. Geneva, 1997.
- [ISO98] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, **ISO 9241 Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 11: Usability principles**. Geneva, 1998.
- [ISO99a] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, **ISO9241 Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) - Part 16: Direct Manipulation Dialogues**. Geneva, 1999.
- [ISO99b] INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, **Introduction to ISO – What are standards. Documento Eletrônico**. [www.iso.ch/infoe/intro.htm#Who does the work](http://www.iso.ch/infoe/intro.htm#Who%20does%20the%20work).
- [ISWN98] ISWorld NET, **User Satisfaction - Recipient Response to the Use of the Output of na IS. Documento Eletrônico**. theweb.badm.sc.edu/grover/ISWorld/satisf.htm.
- [Ives83] IVES, B.; OLSON, M. H. & BAROUDI, J. J., **The Measurement of User Information Satisfaction, Communications of the ACM**, 26(10), October 1983, pp. 785-793.
- [Ives84] IVES, B. & OLSON, M. H., **User Involvement and MIS Success, Management Science**, 30(5), May 1984, pp. 583-603.

- [Jaco92] JACOB, ROBERT J. K. & SIBERT, LINDA E., The perceptual structure of multidimensional input device selection. In: ACM CHI'92 Conference, 1992. **Proceedings**. Monterey, 1992, pp. 211-218.
- [Jaco98] JACOBSEN, N. E.; HERTZUM, M. & JOHN, B. E., The evaluator effect in usability studies: problem detection and severity judgments. In: Human Factors and Ergonomics Society 42nd Annual Meeting. **Proceedings**. 1998, pp. 1336-1340.
- [Jami98] JAMIESON, G. A. & VICENTE K. J., Implications of a Control-Theoretic Approach to Human-Automation- Plant Interface Design. In: 4th Symposium on Human Interaction with Complex Systems. **Proceedings**. Dayton, Ohio, March 22-25 1998, pp. 90-98.
- [Jaya93] JAYARATNE, T., Quantitative methodology and feminist research. In: M. Hammersley (Ed.), **Social research: Philosophy, politics and practice**, pp. 109-123. London: Sage. 1993.
- [Jeff87] JEFFRIES, R. & ROSENBERG, J. K., Comparing a Form-based and a Language-based User Interface Evaluation for Instructing a Mail Program. In: ACM CHI + GI'87 Conference, 1987. **Proceedings**. Toronto, 1987, pp. 261-266.
- [Jeff91] JEFFRIES, R.; MILLER, J. R.; WHARTON, C. & UYEDA, K. M., User Interface Evaluation in the Real World: A Comparison of Four Techniques. In: ACM CHI'91 Conference, 1991. **Proceedings**. New Orleans, Louisiana, April 27 - May 2, 1991, pp. 119-124.
- [John95a] JOHN, B. E. & PACKER, H., Learning and Using the Cognitive Walkthrough Method: A Case Study Approach. In: ACM CHI'95 Conference on Human Factors in Computing Systems. **Proceedings**. Denver, Colorado, May 7-11 1995, pp. 429-436.
- [John96a] JOHNSON, B., Evaluating Evaluation: Understanding When, Where, Why and How Usability Assessment Methods Work. **Documento eletrônico**. *pecan.srv.cs.cmu.edu/afs/cs/user/hci/www/research/projects/evaluating_evaluation.html*. 1996.
- [John96b] JOHNSON, G. I., The usability checklist approach revisited. In Jordan, P. W.; Thomas, B; Weerdmeester, B. A. and McClelland, I. L. (Eds.), **Usability evaluation in industry**. Taylor & Francis, London, 1996, pp. 179-188.
- [Jone97] JONES, IAN, Mixing Qualitative and Quantitative Methods in Sports Fan Research, **The Qualitative Report**, 3(4), December, 1997. (Versão online disponível em www.nova.edu/ssss/QR/QR3-4/jones.html)
- [Jord96a] JORDAN, P. W.; MCCLELLAND, I. L. & THOMAS, B., Introduction. In Jordan, P. W.; Thomas, B; Weerdmeester, B. A. and McClelland, I. L. (Eds.), **Usability evaluation in industry**. Taylor & Francis, London, 1996, pp. 1-3.
- [Jord96b] JORDAN, P. W.; THOMAS, B. & MCCLELLAND, I. L., Issues for usability evaluation in industry: seminar discussions. In Jordan, P. W.; Thomas, B; Weerdmeester, B. A. and McClelland, I. L. (Eds.), **Usability evaluation in industry**. Taylor & Francis, London, 1996, pp. 237-243.
- [Jord98] JORDAN, PATRICK W., **An Introduction to Usability**. Taylor & Francis, London, 1998.
- [Jura95] JURAN, J. M., **Juran's Quality Handbook**. McGraw-Hill, New York, 1995.
- [Kade98] KADHERBAI, TASNIM, Overcoming Inertia within a Large Organization. In: Lesley Trenner & Joanna Bawa, **The Politics of Usability**. Springer-Verlag, London - England, 1998, pp. 35-48.

- [Kali97] KALIMO, RAIJA & SMITH, MICHAEL J., Psychosocial Approach in Occupational Health. In: Gavriel Salvendy (Ed.), **Handbook of Human Factors and Ergonomics**, pp. 1059-1084. New York: John Wiley & Sons Inc. 1997.
- [Kano89] KANO N., NOBUHIRO S., TAKAHASHI F. & TSUJI S., Attractive Quality and Must be Quality, **Quality**, 14(2), pp. 39-48.
- [Kant92] KANT, K., **Introduction to Computer System Performance Evaluation**. McGraw-Hill, New York, 1992.
- [Kapl88] KAPLAN, B. & DUCHON, D., Combining Qualitative and Quantitative Methods in Information Systems Research: A Case Study, **MIS Quarterly**, 12(4), 1988, pp. 571-587.
- [Kapt93] KAPTELININ, V., Item recognition in Menu Selection: The Effect Of Practice. In: INTERCHI'93-CHI'93 & INTERACT'93 Conference. **Adjunct Proceedings**. Amsterdam, The Netherlands, 1993, pp. 183-184.
- [Kara92] KARAT, C.-M.; CAMPBELL, R. & FIEGEL, T., Comparison of Empirical Testing and Walkthrough Methods in User Interface Evaluation. In: ACM CHI'92 Conference, 1992. **Proceedings**. Monterey, 1992, pp. 397-404.
- [Kara93] KARAT, C.-M., Usability Engineering in Dollars and Cents, **IEEE Software**, 26(5), May 1993, pp. 88-89.
- [Kara94a] KARAT, C.-M., A Comparison of User Interface Evaluation Methods. In: Jakob Nielsen and Robert L. Mack, **Usability Inspection Methods**. John Wiley & Sons, Inc. New York, 1994, pp. 203-233.
- [Kara94b] KARAT, C.-M., A Business Case Approach to Usability Cost Justification. In: R. G. Bias & D. J. Mayhew (Eds.), **Cost-Justifying Usability**. Academic Press, Chesnut Hill, Massachusetts, 1994, pp. 45-70.
- [Kara98] KARAT, C.-M., Guaranteeing Rights for the User, **Communications of the ACM**, 41(12), December 1998, pp. 29-31.
- [Kein99] KEINONEN, TURKKA, Usability of Artifacts. **Documento Eletrônico**. www.uiah.fi/projekti/metodi/158.htm. (Originalmente publicado como Chapter 2 - Dimensions of usability - em KEINONEN, TURKKA, **One-dimensional usability - influence of usability on consumers' product preference**, UIAH publication A21. Helsinki, 1998)
- [Keys90] KEYSON, D. K. & PARSONS, K. C., Designing the user interface using interfaces, **Behaviour & Information Technology**, 10(6), 1990, pp. 443-457.
- [Kemp98] KEMP, ELIZABETH A. & KEMP, RAYMOND H., Multimedia Evaluation: Lessons from Learners. Australasian Computer Human Interaction Conference OZCHI'98. **Proceedings**. Adelaide, S. Australia, November 29-December 12 1998, p. 337.
- [Kenn89] KENNEDY, S., Using video in the BNR usability lab, **SIGCHI Bulletin**, 21(2), October 1989, pp. 92-95.
- [Kier85] KIERAS, D. E. & POLSON, P. G., An approach to the formal analysis of user complexity, **International Journal of Man-Machine Studies**, 22, 1985, pp. 365-394.
- [Kier88] KIERAS, D. E., Towards a practical GOMS model methodology for user interface design. In: M. Helander (Ed.), **Handbook of Human-Computer Interaction**. North Holland Elsevier, Amsterdam, 1988, pp. 135-158.

- [Kier92] KIERAS, D. E., Diagrammatic displays for engineered systems: Effects on human performance in interacting with malfunctioning systems, **International Journal of Human-Computer Interaction**, 36, 1992, pp. 861-895.
- [Kill00] KILLINGSWORTH, B. L.; SCHELLENBERGER, R. E. & KLECKLEY, J. W., The Use of Focus Groups in the Design and Development of a National Labor Exchange System, **First Monday**, 5(7), July 2000. (Versão online disponível em firstmonday.org/issues/issue5_7/killingsworth/index.htm).
- [Kim89] KIM, K. K., User Satisfaction: A Synthesis of Three Different Perspectives, **Journal of Information Systems**, 6(4), Fall 1989, pp. 01-12.
- [Kira87] KIRAKOWSKI, J., The Computer User Satisfaction Inventory. In: IEE Colloquium on Evaluation Techniques for Interactive System Design II, 1987. **Proceedings**. London. 1987.
- [Kira88] KIRAKOWSKI, J. & CORBETT, M., Measuring User Satisfaction. In: Jones, D.M. & Winder, R., **People and Computers IV**, Cambridge University Press, UK, 1988.
- [Kira90] KIRAKOWSKI, J. & CORBETT, M., **Effective Methodology for the Study of HCI**. North-Holland, Amsterdam, 1990.
- [Kira93] KIRAKOWSKI, J. & CORBETT, M., SUMI: The software usability measurement inventory, **British Journal of Educational Technology**, 24(3), 1993, pp. 210-212.
- [Kira94] KIRAKOWSKI, J., The Use of Questionnaire Methods for Usability Assessment. **Documento eletrônico**. www.ucc.ie/hfrg/questionnaires/sumi/sumipapp.htm.
- [Kira96] KIRAKOWSKI, J., The software usability measurement inventory: background and usage. In Jordan, P. W.; Thomas, B; Weerdmeester, B. A. and McClelland, I. L. (Eds.), **Usability evaluation in industry**. Taylor & Francis, London, 1996, pp. 169-177.
- [Kiss95] KISSEL, GEORGE V., The Effect of Computer Experience on Subjective and Objective Software Usability Measures. In: ACM CHI'95 Conference on Human Factors in Computing Systems. **Proceedings**. Denver, Colorado, May 7-11 1995, pp. 284-285.
- [Kitc86] KITCHENHAM, B. A. & WALKER, J. G., The meaning of quality. In: Conference on Software Engineering, 1986. **Proceedings**. Southampton - UK, 1986.
- [Kitc96] KITCHENHAM, B. & PFLIEGER, S. H., Software Quality: The Elusive Target, **IEEE Software**, 29(1), January 1996, pp. 12-21.
- [Kivi00] KIVIMÄKI, M.; FELDT, T.; VAHTERA, J. & NURMI, J-E., Sense of Coherence and Health: Evidence from Two Cross-lagged Longitudinal Samples, **Social Science and Medicine**, 50, pp. 583-597, 2000.
- [Krec96] KRECHMER, KEN, Technical Standards: Foundation of the Future, **StandardView**, 4(1), March 1996, pp. 4-8.
- [Krol71] KROLAK, P.; FELTS, W. & MARBLE, G., A Man-Machine Approach Toward Solving the Traveling Salesman Problem, **Communications of the ACM**, 14(5), May 1971, pp. 327-334.
- [Krue83] KRUESS, ELIZABETH, The Human Engineering Task Area, **IEEE Computer**, 16(11), November 1983, pp. 86-93.
- [Krue94] KRUEGER, R. A., **Focus Groups: A Practical Guide for Applied Research**. Sage, Thousand Oaks, CA, 1994, 2nd edition.

- [Kurz96] KURZE, MARTIN, TDraw: a computer-based tactile drawing tool for blind people. In: 2nd International ACM Conference on Assistive Technologies. **Proceedings**. Vancouver, Canada, April 11-12 1996, , pp. 131-138.
- [Kurz98] KURZE, MARTIN, TGuide: a guidance system for tactile image exploration. In: 3rd International ACM Conference on Assistive Technologies. **Proceedings**. Marina del Rey, CA, April 15-17 1998, pp. 85-91.
- [Kust97] KUSTERS, R. J.; SOLINGEN, R. V. & TRIENEKENS, J. J. M., User-perceptions of Embedded Software Quality. **Documento eletrônico**. www.tm.tue.nl/vakgrft/mworkers/rso/paper8.pdf. (Também publicado in STEP97 Conference, 1997. **Proceedings**. IEEE Computer Society Press, ISBN 0-8186-7840-2, July 1997.)
- [Kuva94] KUVAJA, P.; SIMILA, J.; KRZANIK, L.; BICEGO, A.; KOCH, G. & SAUKKONEN, S., **The BOOTSTRAP approach**. Blackwell Business, Oxford UK and Cambridge MA, 1994.
- [Lain98] LAINER, RÜDIGER & WAGNER, INA, Connecting qualities of social use with spatial qualities. Cooperative buildings - Integrating information, organization, and architecture. In: 1st International Workshop on Cooperative Buildings - CoBuild'98. **Proceedings**. Heidelberg, 1998, pp. 191-203.
- [Laws78] LAWSON, H. W.; BERTRAN, M. & SANAGUSTIN, J., The Formal Definition of Human/Machine Communications, **Software - Practice and Experience**, January/February 1978, pp. 51-58.
- [Lea88] LEA, MARTIN, Evaluating User Interface Designs. In: Tony Rubin (Ed.), **User Interface Design for Computer Systems**. Ellis Horwood Limited, Chichester - England, 1984, pp. 134-167.
- [Lece98] LECEROF, A. & PATERNÒ, F., Automatic Support for Usability Evaluation, **IEEE Transactions on Software Engineering**, 24(10), October 1998, pp. 863-888.
- [Lee91] LEE, A. S., Integrating Positivist and Interpretive Approaches to Organizational Research, **Organization Science**, (2), 1991, pp. 342-365.
- [Leha00] LEHANE, P.; TOLEMAN, M. & BENECKE, J., Applying Ecological Interface Design to Experimental Apparatus Used to Monitor a Refrigeration Plant. In: 1st Australasian User Interface Conference AUIC 2000. **Proceedings**. Canberra, Australia, 31 January - 3 February 2000, pp. 41-48.
- [Leit99] LEITE, JAIR C. Introdução à Engenharia de Software. **Documento Eletrônico**. www.dimap.ufrn.br/~jair/ES/c1.htm#sistemas.
- [Levi96a] LEVI, M. D. & CONRAD, F. G., Usability Testing of World Wide Web Sites. **Documento Eletrônico**. stats.bls.gov/ore/pdf/st960150.pdf.
- [Levi96b] LEVI, M. D. & CONRAD, F. G., A Heuristic Evaluation of World Wide Web Prototype, **Interactions**, 3(4), July-August 1996, pp. 50-61.
- [Lewi82] LEWIS, C., Using the "thinking-aloud" method in cognitive interface design. **Research Report RC9265**, IBM T. J. Watson Research Center, Yorktown Heights, NY.
- [Lewi90] LEWIS, C.; POLSON, P.; WHARTON, C. & RIEMAN, J., Testing a Walkthrough Methodology for Theory-Based Design of Walk-Up-and-Use Interfaces. In: CHI'90 Conference. **Proceedings**. Seattle, 1990, pp. 235-242.
- [Lewi91a] LEWIS, J. R., Psychometric Evaluation of an After-Scenario Questionnaire for Computer Usability Studies: The ASQ, **SIGCHI Bulletin**, 23(1), January 1991, pp. 78-81.

- [Lewi91b] LEWIS, C. & POLSON P.G. Cognitive Walkthroughs: A Method for Theory-Based Evaluation of User Interfaces. In: CHI'91 Conference, 1991. **Proceedings**. New Orleans, Louisiana, April 27 - May 2, 1991.
- [Lewi92] LEWIS, C. & MACK, R., Learning to use a text processing system: evidence from "Thinking Aloud" protocols. In: Human Factors in Computer Systems. **Proceedings**. Gaithersberg, Maryland. March 15-17 1992, pp. 387- 392.
- [Lewi93] LEWIS, J. R., IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. **Technical Report 54.786**. Boca Raton, FL: IBM Corp., 1993. (Documento eletrônico disponível em <http://sites.netscape.net/jrlewisinfl/usabqtr.pdf>)
- [Lim96] LIM, K. H.; BENBASAT, I. & TODD, P. A., An experimental investigation of the interactive effects of interface style, instructions, and task familiarity on user performance, **ACM Transactions on Computer-Human Interaction**, 3(1), March 1996, pp. 1-37.
- [Lim98] LIM, K. & USMA, M., Usability Evaluation in the Field: Lessons from a Case-Study Involving Public Information Kiosks. In: 3rd Asian Pacific Computer & Human Interaction APCHI'98. **Proceedings**. Kangawa, Japan. 15 - 17 July 1998, pp. 70-75.
- [Lind99] LINDEMAN, R. W.; SIBERT, J. L. & HAHN, J. K., Towards usable VR: an empirical study of user interfaces for immersive virtual environments. In: CHI 99 Conference on Human Factors in Computing Systems: the CHI is the limit. **Proceedings**. Pittsburgh, PA, May 15 - 20 1999, pp. 64-71.
- [Liu97] LIU, YILI, Software-Use Interface Design. In: Gavriel Salvendy (Ed.), **Handbook of Human Factors and Ergonomics**, pp. 1689-1724. New York: John Wiley & Sons Inc. 1997.
- Löwg92] LÖWGREN, T. M. & NORDVQUIST, T., Knowledge Based Evaluation as Design Support for Graphical User Interfaces. In: CHI'92 Conference, 1992. **Proceedings**. Monterey, 1992, pp. 181-188.
- [Luca75] LUCAS, H. C., Performance and the Use of an Information System, **Management Science**, 21(8), April 1975, pp. 909-919.
- [Lund91] LUNDELL, J. & NOTESS, M., Human Factors in Software Development: Models, Techniques and Outcomes. In: CHI'91 Conference, 1991. **Proceedings**. New Orleans, Louisiana, April 27 - May 2, 1991, pp. 145-151.
- [MacG92] MACGREGOR, J. N., A comparison of the effects of icons and descriptors in videotex menu retrieval, **International Journal of Man-Machine Studies**, 37, 1992, pp. 767-777.
- [Mack93a] MACK, R. & NIELSEN, J., Usability Inspection Methods: Report on a Workshop Held at ACM CHI'92, Monterey, CA, May 3-4, 1992, **SIGCHI Bulletin**, 25(1), January 1993, pp. 28-33.
- [MacL91] MACLEAN, A.; YOUNG, R.; BELLOTTI, V. & MORAN, T., Questions, Options and Criteria: Elements of Design Space Analysis, **Human-Computer Interaction**, 6(3,4), 201-250.
- [MacI93] MACLEOD, M. & REGGER, R., The Development of DRUM: A Software Tool for Video-assisted Usability Evaluation. **Documento eletrônico**. info.lut.ac.uk/research/husat/inuse/usabilitypapers.html.
- [MacI96] MACLEOD, MILES, Performance Measurement and Ecological Validity. In: P Jordan (Ed.), **Usability Evaluation in Industry**. Taylor and Francis, London, UK, 1996, pp. 227-235.

- [Mads99] MADSEN, K. H., The Diversity of Usability Practices, **Communications of the ACM**, 42(5), May 1999, pp. 61-62.
- [Malo85] MALONE, T.W., Designing organisational interfaces. In: CHI '85: Human Factors in Computing Systems. **Proceedings**. San Francisco, April 1985, edited by L. Borman and B. Curtis. New York: ACM.
- [Mand97] MANDEL, THEO, **Elements of User Interface Design**. John Wiley & Sons, New York, NY, 1997.
- [Marc94] MARCHIONINI, G. & CRANE, G., Evaluating Hypermedia and Learning: Methods and Results from the Perseus Project, **ACM Transactions on Information Systems**, 12(1), January 1994, pp. 5-34.
- [Mark94] MARKUS, M. L. & KEIL, M., If we build it, they will come: Designing information systems that people want to use, **Sloan Management Review**, 35(4), 1994, pp. 11-25.
- [Mart73] MARTIN, J., **Design of Man-Computer Dialogues**. Prentice-Hall, Englewood Cliffs, N.J., 1973.
- [Marx94] MARX, MATT & SCHMANDT, CHRIS, Putting people first: specifying proper names in speech interfaces. In: ACM Symposium on User Interface Software and Technology UIST'94. **Proceedings**. Marina del Rey, CA, November 2 - 4 1994, pp. 29-37.
- [Mayh92] MAYHEW, D. J., **Software User Interface Design**. Prentice Hall, Englewood Cliffs, New Jersey, 1992.
- [Mayh94] MAYHEW, D. J. & BIAS, R. G., Organizational Inhibitors and Facilitators. In: R. G. Bias & D. J. Mayhew (Eds.), **Cost-Justifying Usability**. Academic Press, Chesnut Hill, Massachusetts, 1994, pp. 287-318.
- [Mayh99] MAYHEW, D. J., **The Usability Engineering Lifecycle**. Morgan Kaufmann Publishers Inc., San Francisco, California, 1999.
- [Mayk94] MAYKUT, P., & MOREHOUSE, R., **Beginning qualitative research: A philosophic and practical guide**. Falmer Press, London, 1994.
- [McAt98] McATEER, E., Focus Groups/Interviews. In: Jen Harvey (Ed.), **Evaluation Cookbook**. LTDI - Institute for Computer Based Learning/Heriot-Watt University, Edinburgh, 1998, pp. 38-41.
- [McCa77] McCALL, J. A., RICHARDS, P. K. & WALTERS, G. F., Factors in Software Quality. **RADC Reports vols. I, II e III**, 1977.
- [McCa79] McCALL, F. A., An Introduction to Software Quality Metrics. In: John D. Cooper & Matthew J. Fisher (Eds.), **Software Quality Measurement**. A Petrocelli Book, New York, 1979, pp. 127-142.
- [McCl90] McCLELLAND, I. L. & BRIGHAM, F. R., Marketing ergonomics - how should ergonomics be packaged?, **Ergonomics**, 33(5), May 1990, pp. 519-526.
- [McGr94] MCGRAW, K. L., Knowledge Acquisition and Interface Design, **IEEE Software**, November 1994, pp. 90-92.
- [McGr97] MCGRAW, K. L., Defining and designing the performance-centered interface: moving beyond the user-centered interface Interactions, **Interactions**, 4(2), March-April 1997, pp. 19-26. (Versão online disponível em www.acm.org/pubs/articles/journals/interactions/1997-4-2/p19-mcgraw/p19-mcgraw.pdf)

- [Mcke93] MCKERLIE, D. & MACLEAN, A., QOC in Action: Using Design Rationale to Support Design. In: ACM INTERCHI'93 - CHI'93 & INTERACT'93 Conference, 1993. **Proceedings**. Amsterdam, 1993, p. 519.
- [Medy93] MEDYCKYJ-SCOTT, D., Designing geographical information systems for use. In: Medyckyj-Scott, D. & Hearnshaw, H. M., **Human Factors in Geographic Information Systems**. Behaven Press, London, 1993, pp. 87-100.
- [Meln72] MELNYK, VERA, Man-Machine Interface: Frustration, **Journal of the American Society for Information Science**, 23(5), November/December 1972, pp. 392-401.
- [Micr95] MICROSOFT CORPORATION, **The Windows Interface Guidelines for Software Design : An Application Design Guide**. Microsoft Press. Redmond, 1995.
- [Mill81] MILLER, D. P., The depth/ breadth tradeoff in hierarchical computer menus. In: Human Factors Society 25th Annual Meeting. **Proceedings**. Santa Monica, CA , 1981, pp. 296-300.
- [Mill92] MILLER, J. R. & JEFFRIES, R., Usability Evaluation: Science of Trade-offs, **IEEE Software**, 25(9), September 1992, pp. 97-98 e 102.
- [Mill97] MILLER, RAYMOND, Quality and Risk Management. **Documento eletrônico**. www.baz.com/kjordan/swse625/htm/tp-rm.htm. April, 1997.
- [Mitic89] MITCHELL, J. & SHNEIDERMAN, B., Dynamic versus Static Menus: An Exploratory Comparison, **SIGCHI Bulletin**, 20(4), April 1989, pp. 33-37.
- [Monk93] MONK, A.; WRIGHT, P.; HABER, J. & DAVENPORT, L., **Improving your human computer interface: a practical approach**. Prentice Hall International, Hemel Hempstead, 1993.
- [Mora81b] MORAN, T. P., An Applied Psychology of the User, **ACM Computing Surveys**, 13(1), March 1981, pp. 01-11.
- [Mora97] MORAN, T. P.; CHIU, P. & VAN MELLE, W., Pen-based interaction techniques for organizing material on an electronic whiteboard. In: 10th Annual ACM Symposium on User Interface Software and Technology. **Proceedings**. Banff, Canada, October 14 - 17 1997, pp. 45-54.
- [Morg91] MORGAN, K.; MORRIS, R. L. & GIBBS, S., When does a mouse become a rat? or ... comparing the performance and preferences in direct manipulation and command line environments, **Computing Journal**, 34, pp. 265-271.
- [Morg93] MORGAN, D. L. (Ed.), **Successful Focus Groups: Advancing the State of the Art**. Sage, Newbury Park, CA, 1993.
- [Mork98] MORKES, J. & NIELSEN, J., Applying writing guidelines to Web pages. In: CHI'98 Conference. **Proceedings**. Los Angeles, CA, 18-23 April 1998, pp. 321-322. (Versão *online* disponível em www.useit.com/papers/webwriting/rewriting.html - January 6, 1998).
- [Mosi86] MOSIER, J. N., AND SMITH, S. L. Application of Guidelines for Designing User Interface Software, **Behaviour and Information Technology**, 5(1), January 1986, pp. 39 - 46.
- [Moye94] MOYES, JACKIE, When users do and don't rely on icon shape. CHI '94 Conference Companion on Human Factors in Computing Systems. **Proceedings**. Boston, USA, April 24 - 28 1994, pp. 283-284.

- [Mull95] MULLER, M. J.; MCCLARD, A.; BELL, B.; DOOLEY, S.; MEISKEY, L.; MESKILL, J. A.; SPARKS, R. & TELLAM, D., Validating an Extension to Participatory Heuristic Evaluation: Quality of Work and Quality of Work Life. In: ACM CHI'95 Conference on Human Factors in Computing Systems. **Proceedings**. Denver, Colorado, May 7-11 1995, pp. 115-116.
- [Muño00] MUÑOZ, IVETE K., **Validação do Formalismo TAOS para a Análise da Tarefa no Contexto da Concepção de Interfaces Homem-Computador**. (Dissertação de Mestrado em Informática) - DSC/CCT/UFPB, Campina Grande, Fevereiro, 2000.
- [Nau95] NAU, D., Mixing Methodologies: Can Bimodal Research be a Viable Post-Positivist Tool? **The Qualitative Report**, 2(3), December 1995. (Versão *online* disponível em www.nova.edu/ssss/QR/QR2-3/nau.html).
- [Naya95] NAYAK, N. & MRAZEK, D., Analyzing and Communicating Usability Data: Now that you have the data what do you do? In: ACM CHI'95 Conference, 1995. **Proceedings**. Denver, CO, 7-11 May 1995, p. 468.
- [Neal84] NEAL, A. S. & SIMONS, R. M., Playback: A method for evaluating the usability of software and its documentation, **Communications of the ACM**, 23(1), January 1984, pp. 82-96.
- [Neel90] NEELAMKAVIL, F. & MULLARNEY, O., Separating Graphics from Application in the Design of user Interfaces, **The Computer Journal**, 33(5), October 1990, pp. 437-443.
- [Nels95] NELSON, R. RYAN; WHITENER, ELLEN M. & PHILCOX, HENRY H., The Assessment of End-User Training Needs, **Communications of the ACM**, 38(7), July 1995, pp. 2739.
- [Newm80] NEWMAN, S. & SEGEV, E., Evaluate Your Information System, **Journal of Systems Management**, 31(3), March 1980.
- [Nick90] NICKERSON, R. S. & PEW, R. W., Toward More Compatible Human-Computer Interfaces, **IEEE Spectrum**, 27(7), July 1990, pp. 40-43.
- [Niel89a] NIELSEN, JAKOB, Coordinating User Interfaces for Consistency, **SIGCHI Bulletin**, 20(3), January 1989, pp. 63-65.
- [Niel89b] NIELSEN, J. & MOLICH, R., Teaching User Interface Design Based on Usability Engineering, **SIGCHI Bulletin**, 21(1), July 1989, pp. 56 -62.
- [Niel90] NIELSEN, J. & MOLICH, R., Heuristic Evaluation of User Interfaces. In: ACM CHI'90 Conference, 1990. **Proceedings**. Seattle, 1990, pp. 249-256.
- [Niel91] NIELSEN, JAKOB, Trip Report: Usability Metrics and Methodologies, **SIGCHI Bulletin**, 23(2), April 1991, pp. 37-39.
- [Niel92a] NIELSEN, JAKOB, The Usability Engineering Life Cycle, **IEEE Computer**, 25(3), March 1992, pp. 12-22.
- [Niel92b] NIELSEN, JAKOB, Finding Usability Problems Through Heuristic Evaluation. In: ACM CHI'92 Conference, 1992. **Proceedings**. Monterey, 1992, pp. 259-266.
- [Niel92c] NIELSEN, J.; BUSH, R. M.; DAYTON, J. T.; MEND, N. E.; MULLER, M. J. & ROOT, R. W., Teaching experienced developers to design graphical user interfaces. In: ACM CHI'92 Conference, 1992. **Proceedings**. Monterey, California, May 3-7 1992, pp. 557-564.

- [Niel93a] NIELSEN, J. & LANDAUER, T. K., A mathematical model of the finding of usability problems. In: ACM INTERCHI'93 & INTERACT'93 Conference, 1993. **Proceedings**. Amsterdam, 1993, pp. 206-213.
- [Niel93b] NIELSEN, JAKOB, **Usability Engineering**. Academic Press, Cambridge, MA, 1993.
- [Niel93c] NIELSEN, JAKOB, Is Usability Engineering Really Worth It?, **IEEE Software**, 10(6), November 1993, pp. 90-92.
- [Niel93d] NIELSEN, J. & PHILLIPS, V. L., Estimating the Relative Usability of Two Interfaces: Heuristic, Formal, and Empirical Methods Compared. In: ACM INTERCHI'93 - CHI'93 & INTERACT'93 Conference, 1993. **Proceedings**. Amsterdam, 1993, pp. 214-221.
- [Niel93e] NIELSEN, J. & LEVY, J., Subject users preferences versus objective interface performance methods, **Communications of the ACM**, 36, 1993.
- [Niel94a] NIELSEN, J. & MACK, R. L. (Eds.), **Usability Inspection Methods**. John Wiley & Sons, Inc., New York, 1994.
- [Niel94b] NIELSEN, J. & LEVY, J., Measuring Usability: Preference versus Performance, **Communications of the ACM**, 37(4), April 1994, pp. 66-75.
- [Niel94c] NIELSEN, JAKOB, Enhancing the Explanatory Power of Usability Heuristics. In: ACM CHI'94 Conference, 1994. **Proceedings**. Boston - MA, 1994, pp. 152-158.
- [Niel95] NIELSEN, JAKOB, Applying Discount Usability Engineering, **IEEE Software**, 12(1), January 1995, pp. 98-100.
- [Niel96a] NIELSEN, FRAN, Human Behavior: Another Dimension of Standards Sett, **StandardView**, 4(1), September 1996, pp. 36-41.
- [Niel96b] NIELSEN, JAKOB, Top ten mistakes in Web design. **Documento Eletrônico**. www.useit.com/alertbox/9605.html. (AlertBox, May 1996)
- [Niel96c] NIELSEN, JAKOB, Accessible Design for Users With Disabilities. **Documento Eletrônico**. www.useit.com/alertbox/9610.html. (AlertBox, October 1996)
- [Niel96d] NIELSEN, JAKOB, Marginalia of Web Design. **Documento Eletrônico**. www.useit.com/alertbox/9611.html. (AlertBox, November 1996)
- [Niel96e] NIELSEN, JAKOB, Why Frames Suck (Most of the Time). **Documento Eletrônico**. www.useit.com/alertbox/9612.html. (AlertBox, December 1996)
- [Niel97a] NIELSEN, JAKOB, Usability Testing. In: Gavriel Salvendy (Ed.), **Handbook of Human Factors and Ergonomics**. John Wiley & Sons Inc., New York, 1997, pp. 1617-1633.
- [Niel97b] NIELSEN, JAKOB, The Use and Misuse of Focus Groups, **IEEE Software**, 14(1), January 1997, pp. 94-95.
- [Niel97c] NIELSEN, JAKOB, Heuristic Evaluation. **Documento Eletrônico**. www.useit.com/papers/heuristic.
- [Niel97d] NIELSEN, JAKOB, Need for Speed. **Documento Eletrônico**. www.useit.com/alertbox/9703a.html (AlertBox, March 1 1997)
- [Niel97e] NIELSEN, JAKOB, Be Succinct! (Writing for the Web). **Documento Eletrônico**. www.useit.com/alertbox/9703b.html (AlertBox, March 15 1997)

- [Niel97f] NIELSEN, JAKOB, Effective Use of Style Sheets. Documento Eletrônico. www.useit.com/alertbox/9707a.html (AlertBox, July 1 1997)
- [Niel97g] NIELSEN, JAKOB, Search and You May Find. Documento Eletrônico. www.useit.com/alertbox/9707b.html (AlertBox, July 15 1997)
- [Niel97h] NIELSEN, JAKOB, Changes in Web Usability Since 1994. Documento Eletrônico. www.useit.com/alertbox/9712a.html. (Alertbox, December 1997)
- [Niel98] NIELSEN, CHRISTINA, Testing in the Field. In: 3rd Asian Pacific Computer & Human Interaction APCHI'98. **Proceedings**. Kangawa, Japan. 15 - 17 July 1998, pp. 285-290.
- [Niel00] NIELSEN, JAKOB, **Designing Web Usability: The Practice of Simplicity**. New Riders Publishing, Indianapolis - Indiana, 2000.
- [Niga00] NIGAM, AVISHEK, **Usabilidade de Produtos de Suporte ao Desenvolvimento de Aplicações de Banco de Dados**. (Proposta de Dissertação de Mestrado em Engenharia Elétrica) - DEE/CCT/UFPB, Campina Grande, Julho, 2000.
- [NIT00] NOLDUS INFORMATION TECHNOLOGY, The Observer. Documento Eletrônico. www.noldus.com/products/index.html.
- [Nodd99] NODDER, C.; WILLIAMS, G. & DUBROW, D., Evaluating the usability of an evolving collaborative product changes in user type, tasks and evaluation methods over time. In: International ACM SIGGROUP'99 Conference on Supporting Group Work. **Proceedings**. Phoenix, AZ, November 14 - 17 1999, pp. 150-159.
- [Norc89] NORCIO, A. F. & STANLEY, J., Adaptive Human-Computer Interfaces: A Literature Survey and Perspective, **IEEE Transactions on Software Engineering**, 19(2), March/April 1989, pp. 399-408.
- [Norm86] NORMAN, D. A., Cognitive Engineering. In: Norman, D. A. & Draper, S. W. (Eds.), **User Centered System Design: New Perspectives on Human-Computer Interaction**. Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1986, pp. 31-65.
- [Novi99] NOVICK, DAVID G., Using the Cognitive Walkthrough for Operating Procedures, **Interactions**, 6(3), May-June 1999, pp. 31-37.
- [NRC95] Standards, Conformity Assessment, and Trade - Into the 21st Century. Documento Eletrônico. <http://www.nap.edu/readingroom/books/stand/index.html>.
- [Ogaw95] OGAWA, K. & UENO, K., GuideBook - Design Guidelines Database for Assisting the Interface Design Task, **SIGCHI Bulletin**, 27(2), April 1995, pp. 38-39.
- [Ohls78] OHLSON, MARK, System Design Considerations for Graphics Input Devices, **IEEE Computer**, 11 (11), November 1978, pp. 9-18.
- [Olse88] OLSEN, D. R. & HALVERSEN, B. W., Interface usage measurements in a user interface management system. In: ACM UIST'88 1st Symposium User Interface Software and Technology, 1988. **Proceedings**. Banff, Canada, 17-19 October 1988, pp. 102-108.
- [Olso94] OLSON, HOPE, Quantitative "Versus" Qualitative Research: The Wrong Question. Documento eletrônico. www.ualberta.ca/dept/slis/cais/olson.htm.

- [O'Mal84] O'MALLEY, C. E.; DRAPER, S. W. & RILEY, M. S., Constructive interaction: A method for studying human-computer-human interaction. In: IFIP INTERACT'84 First International Conference on Human-Computer Interaction, 1984. **Proceedings**. London, U.K., 4-7 September 1984, pp. 269-274.
- [Oppe88] OPPERMAN, R.; MURCHNER, B.; PAETAU, M.; PIEPER, M.; SIMM, H. & STELLMACHER, I., **Evaluation von Dialogsystemen. Der Software-Ergonomische Leitfaden EVADIS**. De Gruyten, Berlin, 1988.
- [Oppe92] OPPERMAN, R. & REITERER, H., Der Evaluationsleitfaden EVADIS II, **Ergonomie & Informatik**, 15, pp. 25-29, 1992.
- [OSF90] OPEN SOFTWARE FOUNDATION, **OSF/Motif Style Guide**. Prentice Hall, Englewood Cliffs - NJ, 1990.
- [Palm92] PALMITER, SUSAN & ELKERTON, JAY, An Evaluation of Animation Demonstrations for Learning Computer-Based Tasks. In: ACM CHI'92 Conference, 1992. **Proceedings**. Monterey, California, May 3-7 1992, pp. 257-263.
- [Park93] PARKER, N., Geographical Information System Hardware Design. In: Medyckyj-Scott, D. & Hearnshaw, H. M., **Human Factors in Geographic Information Systems**. Behaven Press, London, 1993, pp. 173-184.
- [Paul95] PAULK, M. C.; WEBER, C. V.; CURTIS, B. & CHRISSIS, M. B., **The Capability Maturity Model: Guidelines for Improving the Software Process**. Addison-Wesley, 1995.
- [Payn86] PAYNE, S. J. & GREEN, T. R. G., Task-Action Grammars: A Model of the Mental Representation of Task Languages, **Human-Computer Interaction**, 2, 1986, pp. 91-133.
- [PCMO98] PC MAGAZINE ONLINE, You Can Get There From Here: Eight Ways to Stay in Touch. **Documento eletrônico**. www.zdnet.com/pcmag/features/remotecomtrol/index.html.
- [Pedd92] PEDDIE, JON, **Graphical User Interfaces and Graphic Standards**. McGraw-Hill, Inc., New York, 1st printing, 1992.
- [Pent00] PENTLAND, ALEX, Perceptual Intelligence, **Communications of the ACM**, 43(3), March 2000, pp. 35-44.
- [Penn80] PENNIMAN, W. D. & DOMINICK, W. D., Monitoring and Evaluation of On-Line Information Systems, **Information Processing and Management**, 16(1), 1980, pp. 17-35.
- [Pern95] PERNICE, K. & BUTLER, M. B., Database Support for Usability Testing, **Interactions**, 2(1), January 1995, pp. 11-20.
- [Pfle91] PFLEEGER, S. L., **Software Engineering: the Production of Quality Software**. Macmillan, 2nd edition, 1991.
- [Pica00] PICARD, R. W., **Affective Computing**. The MIT Press, 1st paperback edition, 2000.
- [Ping89] PINGARD, P. & NANARD, M., Génération d'Interface Dynamiquement Adaptable, **Rapport Technique Interne**, Centre de Recherche en Informatique de Montpellier. Montpellier, Mai 1989.
- [PolI90] POLLIER, AGNÈS, Étude Descriptive d'une Activité de Conception-Evaluation Ergonomique d'un Système de Messagerie Vocale. **Rapport de Recherche N°1160 Programme 8**, INRIA. Le Chesnay, Février, 1990.

- [Pols88] POLSON, P. G., The Consequences of Consistent and Inconsistent User Interfaces. In: R. Guindon (Ed.), **Cognitive Science and Its Applications for Human-Computer Interaction**. Lawrence Erlbaum Associates, Hillsdale, NJ, 1988, pp. 59-108.
- [Pols90] POLSON, P.G. & LEWIS, C., Theory-based design for easily learned interfaces, **Human-Computer Interaction**, 5, 1990, pp.191-220.
- [Pols92] POLSON, P., LEWIS, C., RIEMAN, J. & WHARTON, C., Cognitive walkthroughs: A method for theory-based evaluation of user interfaces, **International Journal of Man-Machine Studies**, 36, 1992, pp. 741-773.
- [Popo95] POPOWICZ, ALISON, Collecting User Information on a Limited Budget, **SIGCHI Bulletin**, 27(4), October 1995, pp. 23-28.
- [Port97] PORTIGAL, STEVE, Visual Interaction Design: Design as a Cultural Activity, **SIGCHI Bulletin**, 29(3), July 1997. (Versão *online* disponível em www.acm.org/sigchi/bulletin/1997.3/vid.html)
- [Powe00] POWELL, THOMAS A., **Web Design: The Complete Reference**. Osborne/McGraw-Hill, Berkeley - CA, 2000.
- [Powe73] POWERS, R. F. & DICKSON, G. W., MIS Project Management: Myths, Opinions, and Reality, **California Management Revue**, 15(3), Spring 1973, pp. 147-156.
- [Pree90] PREECE, J. & KELLER, L., **Human-Computer Interaction**, Prentice Hall International (UK) Ltd. - Hertfordshire, 1990.
- [Pree94] PREECE, J., ROGERS, Y., SHARP, H., BENYON, D., HOLLAND, S. & CAREY, T., **Human-Computer Interaction**. Addison-Wesley Publishing Company, Wokingham, 1994.
- [Pree95] PREECE, JENNY, **A Guide to Usability: Human Factors in Computing**, Addison-Wesley - Reading, MA. 1995.
- [Pres95] PRESSMAN, R. S., **Engenharia de Software**. Makron Books, São Paulo, 1995.
- [Pres99] PRESTON, ALICE, Remote Usability Testing Tools, **Usability Interface**, 5(3), January 1999. (Versão *online* disponível em stc.org/pics/usability/newsletter/9901-remote-tools.html)
- [Prüm91] PRÜMPER, J., FRESE, M.; ZAPF, D. & BRODBECK, F. C., Errors in Computerized Office Work: Differences Between Novice and Expert Users, **SIGCHI Bulletin**, 23(2), April 1991, pp. 63-66.
- [Quei94] QUEIROZ, JOSÉ EUSTÁQUIO RANGEL DE, **Validação de uma Metodologia de Avaliação de Projetos de Interfaces Usuário-Computador**. (Dissertação de Mestrado em Engenharia Elétrica) - DEE/CCT/UFPB, Campina Grande, Julho, 1994.
- [Quei96] QUEIROZ, J. E. R. DE & TURNELL, M. DE F. Q. V., Avaliação de Projetos de Interfaces Usuário-Computador: Proposição de uma Estratégia Qualitativa Adaptativa. **Relatório Técnico RT00098 COPELE/CCT/UFPB**. Campina Grande, Junho, 1996.
- [Quei97a] QUEIROZ, J. E. R. DE & TURNELL, M. F. Q. V., Proposição de um Espaço de Trabalho Dedicado à Avaliação de Processos Interativos Usuário-Computador. **Relatório Técnico RT00118 COPELE/CCT/UFPB**. Campina Grande, Abril, 1997.

- [Quei97b] QUEIROZ, J. E. R. DE & TURNELL, M. F. Q. V., Proposta de Abordagem para Seleção de Estratégias Avaliatórias de Processos Interativos. In: Congresso Latinoamericano de Ingeniería de Software, 1997. **Anales**. Valparaíso, Chile, Noviembre 3-7, 1997, pp. 397-404.
- [Quei97c] QUEIROZ, J. E. R. DE & TURNELL, M. DE F. Q. V., Proposição de Linhas de Ação para o Laboratório de Avaliação de Processos Interativos Usuário-Computador - LAPI do DEE/UFPB. **Relatório Técnico COPELE/CCT/UFPB**. Campina Grande, Agosto, 1997.
- [Quei98a] QUEIROZ, J. E. R. DE & TURNELL, M. F. Q. V., Avaliando a Avaliação: Um SIG como Estudo de Caso. **Relatório Técnico RT00170 COPELE/CCT/UFPB**. Campina Grande, Julho, 1998.
- [Quei98b] QUEIROZ, J. E. R. DE & TURNELL, M. F. Q. V., Evaluating the Quality of Human-Computer Interfaces According to Specific Contexts. In: International Conference on Systems, Man and Cybernetics - SMC'98, 1998. **Proceedings**. Hyatt La Jolla, San Diego, CA, October 11-14, 1998, pp. 1296-1301.
- [RACE95] RACE, The Guidance Project (RACE 1067). **Documento eletrônico**. www.mrc-cbu.cam.ac.uk/amodeus/guidance/.
- [Ragi87] RAGIN, C. C., **The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies**. University of California Press, Berkeley/London, 1987.
- [Rask00] RASKIN, JEF, **The Humane Interface**. ACM Press (New York - NY)/Addison-Wesley Longman Inc. (Reading - MA), 2000.
- [Reas81] REASON, P., & ROWAN, J., **Human inquiry: A sourcebook of new paradigm research**. John Wiley and Sons, Chichester, 1981.
- [Rehe96] REHESAAR, HUGO, International Standards: Practical or Just Theoretical?, **StandardView**, 4(3), September 1996, pp. 123-127.
- [Reis81] REISNER, P., Formal Grammar and Human Factors Design of an Interactive Graphics System, **IEEE Transactions on Software Engineering**, SE-7, 1981, pp. 229-240.
- [Reit95] REITERER, HARALD, IDA - User Interface Design Assistance, **SIGCHI Bulletin**, 27(2), April 1995, pp. 46-47.
- [Reki96] REKIMOTO, JUN, "Tilting Operations for Small Screen Interfaces. In: ACM Symposium on User Interface Software and Technology UIST'96. **Proceedings**. Seattle, WA, November 6 - 8 1996, pp. 167-168.
- [Resn95] RESNICK, PAUL & VIRZI, ROBERT A., Relief from the audio interface blues: expanding the spectrum of menu, list, and form styles, **ACM Transactions on Computer-Human Interaction**, 2(2), June 1995, pp. 145-176.
- [Rett92] RETTIG, MARC, Interface Design When You Don't Know How, **Communications of the ACM**, 35(1), January 1992, pp. 29-34.
- [Ride91] RIDEOUT, TOM, Changing Your Methods From The Inside, **IEEE Software**, 24(5), May 1991, pp. 99-100 e 111.
- [Riem91] RIEMAN, J.; DAVIES, S.; HAIR, D. C.; ESEMPARE, M.; POLSON, P.; LEWIS, L., An Automated Cognitive Walkthrough. In: ACM CHI'91 Conference, 1991. **Proceedings**. New Orleans, Louisiana, April 27 - May 2, 1991, pp. 427-428.

- [Riem95] RIEMAN, J.; DAVIES, S.; HAIR, D. C.; ESEMPARE, M.; POLSON, P.; LEWIS, L., An Automated Cognitive Walkthrough. In: ACM CHI'95 Conference, 1995. **Proceedings**. Denver, CO, 7-11 May 1995, pp. 387-388.
- [Rizz97] RIZZO, A.; MARCHIGIANI, E. & ANDREADIS A., The AVANTI project: prototyping and evaluation with a cognitive walkthrough based on the Norman's model of action. In: ACM Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques DIS'97. **Proceedings**. Amsterdam, August 18 - 20, 1997, pp. 305-309.
- [Robe66] ROBERTS, L. G., A Graphical Service System With Variable Syntax Communication, **Communications of the ACM**, 9(3), March 1966, pp. 173-176.
- [Robe83] ROBERTS, T. L. & MORAN, T. P., The Evaluation of Text Editors: Methodology and Empirical Results, **Communications of the ACM**, 26(3), April 1983, pp. 265-283.
- [Rodr98] RODRIGUES, C. A. G., O Constante Aprendizado do Usuário, **PC Magazine Brasil**, 8(10), Outubro de 1998, pp. 54-55.
- [Roge92] ROGERS, S., VERECZEKEI, M., RICHIER, A. & WRIGHT, M., **ESPRIT Advanced Business and Home Systems - Peripherals: The Synopses**. Luxembourg: Office of Official Publications of the European Communities. 1992.
- [Rohn98] ROHN, JANICE A., Creating Usable E-Commerce Sites, **StandardView**, 6(3), September 1998, pp. 110-115.
- [Root83] ROOT, R. W. & DRAPER, S., Questionnaires as a Software Evaluation Tool. In: ACM CHI'83 Conference, 1983. **Proceedings**. Boston - MA, 12-15 December 1983, pp. 83-87.
- [Root93] ROOT, R. W. & UYEDA, K. M., A Headsup on Guidestyles: Report on the CHI'92 Styleguide SIG, **SIGCHI Bulletin**, 25(3), July 1993, pp. 32-35.
- [Rose96] ROSENZWEIG, ELIZABETH, Design Guidelines for Software Products: A Common Look and Fell or a Fantasy?, **Interactions**, 3(6), September-October 1996, pp. 21-26.
- [Ross97] ROSS, FIONN, Key Issues in Preparing for Design. **Documento eletrônico**. www.geocities.com/SiliconValley/Hills/5776/chapter1.html. 1997.
- [Rous97] ROUSE, W. B. & BOFF, K. R., Assessing Cost/Benefits of Human Factors. In: Gavriel Salvendy (Ed.), **Handbook of Human Factors and Ergonomics**. John Wiley & Sons Inc., New York, 1997, pp. 1617-1633.
- [Rowl92] ROWLEY, D. E. & RHOADES, D. G., The cognitive jogthrough: a fast-paced user interface evaluation procedure. In: ACM Conference on Human Factors in Computing Systems CHI'92. **Proceedings**. Monterey, California, 3 - 7 May 1992, pp. 389 - 395.
- [Rowl94] ROWLEY, D. E., Usability Testing in the Field: Bringing the Laboratory to the User. In: ACM CHI'94 Conference, 1994. **Proceedings**. Boston - MA, 1994, pp. 252-257.
- [Roy90] ROY, RANJIT, **A Primer on the Taguchi Method**, Van Nostrand Reinhold, New York, 1990.
- [Rubi94] RUBIN, JEFFREY, **Handbook of Usability Testing**, John Wiley & Sons, Inc., New York, 1994.
- [Salv97] SALVENDRY, GAVRIEL (Ed.), **Handbook of Human Factors and Ergonomics**. John Wiley & Sons Inc., New York, 2nd edition, 1997.

- [Salv98] SALVADOR, T. & HOWELLS, K., Focus Troupe: Using Drama to Create Common Context for New Product Concept End-User Evaluations. In: CHI'98 Conference. **Proceedings**. Los Angeles, CA, 18-23 April 1998, pp. 251-252.
- [Sato99] SATO, S. & SALVADOR, T., Playacting and Focus Troupes: Theater techniques for creating quick, intense, immersive, and engaging focus group sessions, **Interactions**, 6(5), Setembro/Outubro 1999, pp. 35-41.
- [Savi96] SAVIDIS A.; STEPHANIDIS, C.; KORTE, A.; CRISPIEN, K. & FELLBAUM, K., A generic direct-manipulation 3D-auditory environment for hierarchical navigation in non-visual interaction. In: 2nd Annual ACM Conference on Assistive Technologies. **Proceedings**. Vancouver, Canada, April 11 - 12, 1996, pp. 117-123.
- [Scap88] SCAPIN, D. L., Vers des Outils Formels de Description des Tâches Orientées Conception d'Interfaces. **Rapport de Recherche N° 893**. Le Chesnay, Septembre, 1988.
- [Scap89] SCAPIN, D. L. & PIERRET-GOLBREICH, Towards a Method for Task Description: MAD. **Rapport de Recherche**. INRIA, Rocquencourt, France, Août, 1989.
- [Scha90] SCHACH, S. R., **Software Engineering**. Richard D. Irwin, Inc. & Aksen Associates, 1990.
- [Scha93] SCHACH, S. R., **Software Engineering**. Richard D. Irwin, Inc. & Aksen Associates, 2nd edition, 1993.
- [Scho98] SCHOLTZ, J. & LASKOWSKI, S. & DOWNEY, L., Developing Usability Tools and Techniques for Designing and Testing Web Sites. In: 4th Conference on Human Factors and the Web. Basking Ridge, New Jersey, Friday, June 5, 1998. (Versão *online* disponível em www.research.att.com/conf/hfweb/proceedings/scholtz/index.html)
- [Scho99] SCHOLTZ, J., A Case Study: Developing a Remote, Rapid, and Automated Usability Testing Methodology for On-Line Books. In: 32th Annual Hawaii International Conference on System Sciences HICSS'99. Maui, Hawaii, 5 - 8 January 1999. (Versão *online* disponível em dlib.computer.org/conferen/hicss/0001/pdf/00012015.pdf)
- [Schr92] SCHREIER, J. R., Reducing Stress Associated with Participating in a Usability Test. In: Human Factors Society 36th Annual Meeting, 1992. **Proceedings**. 1993, pp. 1210-1214.
- [Scri67] SCRIVEN, M., The methodology of evaluation. in: R. Tyler, R. Gagne and M. Scriven (Eds.), **Perspectives of Curriculum Evaluation**. Chicago: Rand McNally, 1967.
- [Sear98] SEARS, A. & HESS, D. J., The Effect of Task Description Detail on Evaluator Performance with Cognitive Walkthroughs. In: ACM CHI'98 Conference, 1998. **Proceedings**. Los Angeles, CA, 18-23 April 1998, pp. 259-260.
- [Sebi94] SEBILLOTE, S., La Conception de Scénarios Interactifs: Analyse de l'Activité. **Rapport de Recherche**. INRIA, Rocquencourt, France, Novembre, 1994.
- [Sena87] SENAY, H. & STABLER, E. P., Online help system usage: An empirical investigation. In: 2th International Conference on Human-Computer Interaction, 1987. **Abridged Proceedings**. Honolulu, HI, 10-14 August 1987, p. 244.
- [Shac71] SHACKEL, B., Human factors in P.L.A. meat handling automation scheme: A case study and some conclusions, **International Journal of Productivity Research**, 9(1), 1971, pp. 95-121.

- [Shac84] SHACKEL, B., The concept of usability. In: J. Bennett, D. Case, J. Sandelin & M. Smith (Eds.), **Visual Display terminals**. Prentice-Hall, Englewood Cliffs - NJ, 1984, pp. 45-87.
- [Shac86] SHACKEL, B., Ergonomics in design for usability. In: M.D. Harrison & A.F. Monk (Eds.), **People and Computers: Designing for Usability**. HCI '86. **Proceedings**. Cambridge University Press, Cambridge, 1986.
- [Shac90] SHACKEL, B., Human factors and usability. In J. Preece & L. Keller (Eds.), **Human-Computer-Interaction: Selected Readings**. Hemel Hempstead, Prentice-Hall. 1990.
- [Shac91] SHACKEL, B., Usability: Context, Framework, Definition, Design and Evaluation. In: B. Shackel and S. Richardson (Eds.), **Human Factors for Informatics Usability**. Cambridge University Press, Cambridge, 1991.
- [Sher97] SHERIDAN, T. B., Supervisory Control. In: Gavriel Salvendy (Ed.), **Handbook of Human Factors and Ergonomics**, pp. 1295-1327. New York, John Wiley & Sons Inc. 1997.
- [Shne79] SHNEIDERMAN, BEN, Human Factors Experiments in Designing Interactive Systems, **IEEE Computer**, 12(12), December 1979, pp. 9-19.
- [Shne82] SHNEIDERMAN, BEN, System Message Design: Guidelines and Experimental Results. In: Albert Badre and Ben Shneiderman (Eds.), **Directions in Human-Computer Interface**. Norwood, 1982, pp. 55-78.
- [Shne84] SHNEIDERMAN, BEN, The Future of Interactive Systems and the Emergence of Direct Manipulation. In: Yannis Vassiliou (Ed.), **Human Factors and Interactive Computer Systems**. Norwood, 1984, pp. 01-27.
- [Shne87] SHNEIDERMAN, BEN, **Designing the User Interface: Strategies for Effective Human-Computer Interaction**, Addison-Wesley Publishing Company - Reading, 1987.
- [Shne89] SHNEIDERMAN, B. & KEARSLEY, G., **Hypertext Hands-On! An Introduction to a New Way of Organizing and Accessing Information**. Addison-Wesley, Reading, MA, 1989.
- [Shne93] SHNEIDERMAN, BEN, **Sparks in Innovation in Human-Computer Interaction**. Ablex Publishing, 1993.
- [Shne98] SHNEIDERMAN, B.; NIELSEN, J.; BUTLER, S. & LEVI, M., Is the Web really different from everything else? In: CHI'98 Conference. **Proceedings**. Los Angeles, CA, 18-23 April 1998, pp. 92-93.
- [Shne00] SHNEIDERMAN, BEN, Universal Usability, **Communications of the ACM**, 43(5), May 2000, pp. 85-91.
- [Sibe00] SIBERT, LINDA E. & JACOB, ROBERT J. K., Evaluation of Eye Gaze Interaction. In: CHI 2000 Conference on Human Factors in Computing Systems. **Proceedings**. The Hague, Netherlands, April 1 - 6 2000, pp. 281-288.
- [Sieg97] SIEGEL, J. & BAUER, M., A Field Usability Evaluation of a Wearable System. 1st International Symposium on Wearable Computers (ISWC '97). **Proceedings**. Cambridge, Massachusetts. October 13-14 1997, pp. 18-23.
- [Silv92] SILVA, JOSÉ B. F., **Avaliação Qualitativa de Interfaces Bancárias com Proposta de Diretrizes de Projeto**. (Dissertação de Mestrado em Informática) - DCC/CCT/UFPB, Campina Grande, Julho, 1992.

- [Sioc89] SIOCHI, A. C. & HARTSON, H. R., Task-oriented representation of asynchronous user interfaces. In: CHI'89 Conference on Human Factors in Computing Systems. **Proceedings**. Austin, Texas, April 30-May 4 1989, pp. 183-188.
- [Sioc91] SIOCHI, A. C. & EHRICH, R. W., Computer Analysis of User Interfaces Based on Repetition in Transcripts of User Sessions, **ACM Transactions on Information Systems**, 9(4), October 1991, pp. 309-335.
- [Sioc94] SIOCHI, A. C. & HIX, D., A Study of Computer-supported User Interface Evaluation Using Maximal Repeating Pattern Analysis. In: ACM CHI'94 Conference, 1994. **Proceedings**. Boston, Massachusetts, 1994, pp. 301-305.
- [Slau94] SLAUGHTER, L. A., HARPER, B. D. & NORMAN, K. L., Assessing the Equivalence of Paper and On-line versions of the QUIS 5.5. In: 2nd Annual Mid-Atlantic Human Factors Conference. **Proceedings**. Washington, D.C., 1994, pp. 87-91. (Versão *online* disponível em lap.umd.edu/quis/publications/slaughter1994.html)
- [Smit86a] SMITH, S. L. & MOSIER, J. N., Standards versus Guidelines for Designing User Interface Software, **Behaviour and Information Technology**, 5(1), January 1986, pp. 47-61.
- [Smit86b] SMITH, S. L. & MOSIER, J. N., Guidelines for designing User Interface Software. **Report ESD-TR-86-278 MTR-10090**. The MITRE Corporation, Bedford, MA, 1986.
- [Smoo95] SMOOT, OLIVER R., Tension and synergism between standards and intellectual property, **StandardView**, 3(2), June 1995, pp. 60-67.
- [Snee78] SNEERINGER, JAMES, User-interface Design for Text Editing: A case Study, **Software - Practice and Experience**, 8(5), September/October 1978, pp. 543-557.
- [Soli97] SOLINGEN, R. V. & BERGHOUT, E., Improvement by goal-oriented measurement - Bringing the Goal/Question/Metric approach up to Level 5. In: European Software Engineering Process Group Conference (E-SEPG), 1997. **Proceedings**. Amsterdam, The Netherlands. June 16-20, 1997.
- [Soli99] SOLINGEN, R. V.; UIJTREGT, A. V.; KUSTERS, R. & TRIENEKENS, J., Tailoring product focused SPI - Application and customisation of PROFES in Tokheim. **Documento eletrônico**. www.tm.tue.nl/vakgrfit/mworkers/rso/paper11.pdf (Também publicado in 1st PROFES Conference. **Proceedings**. Oulu, Finland. June 22-24 1999, www.ele.vtt.fi/profes99)
- [Somm97] SOMMERVILLE, IAN, **Software Engineering**. Addison-Wesley, Reading, Massachusetts, 1997.
- [Sond82] SONDHEIMER, N. K. & RELLES, N., Human Factors and User Assistance in Interactive Computing Systems: An Introduction, **IEEE Transactions on Systems, Man and Cybernetics**, SMC-12(2), March/April 1982, pp.102-107.
- [Spen85] SPENCE, R. H., **Computer Usability Testing and Evaluation**. Prentice Hall, Englewood Cliffs, N.J., 1985.
- [Spil81] SPILIOTOPOULOS, V. & SHACKEL, B. Towards a computer interview acceptable to the naive user, **International Journal of Man-Machine Studies**, 14, 1981, pp. 77-90.
- [Squi99] SQUIRES, D. & PREECE, J., Predicting Quality in Educational Software: Evaluating for learning, usability and the synergy between them, **Interacting with Computers**, 11 (5), 1999, pp. 467-483.

- [Stan97a] STANNEY, K. M.; MAXEY, J. & SALVENDRY, G., Socially Centered Design. In: Gavriel Salvendry (Ed.), **Handbook of Human Factors and Ergonomics**, pp. 637-656. New York: John Wiley & Sons Inc. 1997.
- [Stan97b] STANTON, N. A. GALE, A. & ASHLEIGH, M. (1997) CAFE OF EVE in the control room. In: Harris D, (Ed) **Engineering Psychology and Cognitive Ergonomics: integration of theory and application**. Avebury Technical: Aldershot.
- [Stew90] STEWART, T., SIOIS - Standard Interfaces or Interface Standards. In: IFIP INTERACT'90 Conference, 1990. **Proceedings**. Cambridge, U. K., 27-31 August 1990, pp. 29-34.
- [Stew98] STEWART, TOM, Standards and Style Guides – a Cross-Cultural Perspective. In: Lesley Trenner & Joanna Bawa, **The Politics of Usability**. Springer-Verlag, London - England, 1998, pp. 157-167.
- [Stew00] STEWART, TOM, Ergonomics user interface standards: are they more trouble than they are worth? **Documento eletrônico**. <http://www.system-concepts.com/stds/standards.pdf>. (Documento originalmente escrito para o HFES2000 e publicado em **Ergonomics**, 43(7), July 2000, pp. 1030-1044).
- [Stro94] STRONG, GARY, New Directions in Human-Computer Interaction Education, Research and Practice. CHI'94 and Snowbird'94 Conferences. **HCI Workshop Report**. (Versão *online* disponível em www.sei.cmu.edu/arpa/hci/directions/.)
- [Sugi98] SUGIURA, ATSUSHI & KOSEKI, YOSHIYUKI, A user interface using fingerprint recognition: holding commands and data objects on fingers. In: 11th Annual ACM Symposium on User Interface Software and Technology. **Proceedings**. San Francisco, CA, November 1 - 4 1998, pp. 71-79.
- [Suhm99] SUHM, B.; MYERS, B. & WAIBEL, A., Model-based and Empirical Evaluation of Multimodal Interactive Error Correction. CHI'99 Conference on Human Factors in Computing Systems: the CHI is the limit. **Proceedings**. Pittsburgh, PA, May 15 - 20 1999, pp. 584-591.
- [Sula96] SULAIMAN, SUZIAH, Usability and The Software Production Life Cycle. **Documento eletrônico**. www.acm.org/sigchi/chi96/Doctor-Consort/sulaiman/ss_txt.htm.
- [Sun99] SUN MICROSYSTEMS INC., **Java Look and Feel Design Guidelines**. Addison-Wesley, 1999.
- [Sven91] SVENDSON, G. B., The influence of interface style on problem solving, **International Journal of Man-Machine Studies**, 35, 1991, pp. 379-397.
- [Swal97] SWALLOW, J.; HAMELUCK, D. & CAREY, T. User Interface Instrumentation for Usability Analysis: A Case Study. In: Cascon'97. Toronto, Ontario, November 1997. (Versão *online* disponível em watserv1.uwaterloo.ca/~tcarey/casestudy.html).
- [Swan74] SWANSON, E. B., Management Information Systems: Appreciation and Involvement, **Management Science**, 21(2), October 1974, pp. 179-188.
- [Swee93] SWEENEY, M.; MAGUIRE, M. & SHAKEL, B., Evaluating User-computer Interaction: A Framework, **International Journal of Man-Machine Studies**, 38(4), 1993, pp. 689-711.
- [Tanr00] TANRIVERDI, VILDAN & JACOB, ROBERT J. K., Interacting with eye movements in virtual environments . In: CHI 2000 Conference on Human Factors in Computing Systems. **Proceedings**. The Hague, Netherlands, April 1 - 6 2000, pp. 265-272.

- [Tell99] TELLIOGLU, HILDA & WAGNER, INA, Cooperative work across cultural boundaries in systems design, **Scandinavian Journal of Information Systems**, 11(1), 1999. (Versão online disponível em http://iris.informatik.gu.se/sjis/magazine/vol11no1/tellioglu_and_wagner.pdf).
- [Terv96] TERVONEN, ILKKA, Support for Quality-based Design and Inspection, **IEEE Software**, 29(1), January 1996, pp. 44-54.
- [Terw96] TERWILLIGER, R. B. & POLSON, P. G., Task elaboration or label following: an empirical study of representation in human-computer interaction. In: Conference Companion on Human Factors in Computing Systems CHI '96: Common Ground. **Proceedings**. Vancouver, Canada, April 13-18 1996, pp. 201-202.
- [Tetz90] TETZLAFF, L. & CLEVELAND, L., Application logging for usability analysis of graphical user interfaces. **Research Report**. IBM Research Division. T.J. Watson Research Center, Yorktown Heights, New York, 1990.
- [Tetz91] TETZLAFF, L. & SCHWARTZ, D. R., The Use of Guidelines of Interface Design. In: CHI'91 Conference, 1991. **Proceedings**. New Orleans, 1991, pp. 329-333.
- [Thim94] THIMBLEBY, HAROLD, Formulating Usability, **SIGCHI Bulletin**, 26(2), April 1994, pp. 59-64.
- [Thim99] THIMBLEBY, HAROLD, A future for Usability Engineering. CS-99-01, School of Computing Science Technical Report Series. ISSN 1462-0871. **Technical Report**. 1999.
- [Thim00] THIMBLEBY, H.; JONES, M. & MARSDEN, G., Empowering Consumers with Usability Certificates. In: BCS HCI Conference 2000. **Proceedings**. in press, 2000. (Versão online disponível em www.cs.mdx.ac.uk/harold/srf/default.htm#r1)
- [Thom94] THOMAS, M. & McGARRY, F., Top-Down versus Bottom-Up Process Improvement, **IEEE Software**, 27(7), July 1994, pp. 12-13.
- [Thom96] THOMAS, B., 'Quick and dirty' usability tests. In Jordan, P. W.; Thomas, B; Weerdmeester, B. A. and McClelland, I. L. (Eds.), **Usability evaluation in industry**. Taylor & Francis, London, 1996, pp. 107-114.
- [Thom00] THOMAS, B. H. & DEMCZUK, V., Evaluation of Animation Effects to Improve Indirect Manipulation. In: 1st Australasian User Interface Conference AUIC 2000. **Proceedings**. Canberra, Australia, 31 January - 3 February 2000, pp. 110-117
- [TMS00] TANTARA MANAGEMENT SERVICES, What is quality -- the experts opinion. **Documento eletrônico**. www.tantara.ab.ca/what_q.htm.
- [Togn89] TOGNAZZINY, BRUCE, Achieving Consistency for the Macintosh. In: Jacob Nielsen (Ed.), **Coordinating User Interfaces for Consistency**. Academic Press, Boston, MA, 1989, pp. 74-88.
- [Togn90] TOGNAZZINI, BRUCE, Consistency. In: Brenda Laurel (Ed.), **The Art of Human-Computer Interface**. Addison-Wesley Publishing Co., Inc. Reading, 1990, pp. 75-77.
- [Tole98] TOLEMAN, MARK A. & TOLEMAN, JESSICA M. , User Experiences and a Usability Inspection of an Electronic Services Environment for Students. Australasian Computer Human Interaction Conference OZCHI'98. **Proceedings**. Adelaide, S. Australia, November 29-December 12 1998, pp. 87-93.

- [Trau92] TRAUTH, E. M. & COLE, E., The Organizational Interface: A Method for Supporting End Users of Packaged Software, *MIS Quarterly*, 16(1), March 1992. (Versão *online* disponível em www.cba.neu.edu/~etrauth/works/misq92.txt)
- [Treu94] SIEGFRIED, TREU, *User Interface Evaluation*. Plenum Press, New York, NY, 1994.
- [Treu98] SIEGFRIED, TREU, Need for Multi-Aspect Measures to Support Evaluation of Complex Human-Computer Interfaces. In: 4th Symposium on Human Interaction with Complex Systems HICS'98. *Proceedings*. Dayton, Ohio, March 22-25 1998, pp. 182-193.
- [Trum93] TRUMBLY, J. E.; ARNETT, K. P. & MARTIN, M. P., Performance effect of matching computer interface characteristics and user skill level, *International Journal of Man-Machine Studies*, 38, pp. 713-724.
- [Tyld88] TYLDESLEY, D. A., Employing usability engineering in development of office products, *Computer Journal*, 31(5), May 1988, pp. 431-436.
- [Vala93] VALACICH, J. S.; MENNECKE, B. E.; WACHTER, R. & WHEELER, B. C., Computer-Mediated Idea Generation: The Effects of Group Size and Group Heterogeneity. 36th Hawaii International Conference on Systems Science, IV, 1993. *Proceedings*. 1993, pp. 152-160.
- [Vand95] VANDERDONCKT, JEAN, SIERRA - An Interactive System for Ergonomic Realization of Applications, *SIGCHI Bulletin*, 27(2), April 1995, pp. 50-51.
- [VanV96] VAN VIANEN, E.; THOMAS, B. & NIEUWKASTEELE, M. VAN, A combined effort in the standardization of user interface testing. In Jordan, P. W.; Thomas, B; Weerdmeester, B. A. and McClelland, I. L. (Eds.), *Usability evaluation in industry*. Taylor & Francis, London, 1996, pp. 1-17.
- [Virv00] VIRVOU, MARIA & TSIRIGA, VICTORIA, Involving Effectively Teachers and Students in the Life Cycle of an Intelligent Tutoring System, *Educational Technology & Society*, 3(3), 2000, pp. 511-521.
- [Virz91] VIRZI, R. A., A preference evaluation of three dialing plans for a residential, phone-based information service. In: Human Factors Society 35th Annual Meeting, 1991. *Proceedings*. San Francisco, CA, 2-6 September, 1991.
- [Visc97] VISCIOLA, M. & SERRA, A., Il Software Design per l'Usabilità e la Qualità del Software. In: VI^o Congresso Nazionale della Società Italiana di Ergonomia: Governo delle Tecnologie, Efficienza e Creatività. *Atti*. Bologna, 15-17 Settembre, Monduzzi Editore, 1997. (Versão *online* em inglês disponível em www.qualitylab.com/customer.html).
- [Vora95] VORA, P. R. & HELANDER, M. G., A Teaching method as an alternative to the concurrent think-aloud method for usability testing. In Y. Anzai, K. Ogawa & H. Mori, *Symbiosis of Human and Artifact*, North-Holland, 1995, pp.375-380.
- [Vora98] VORA, PAWAN R., Design/Methods Tools Designing for the Web: a survey, *Interactions*, 5(3), May-June 1998, pp. 13-30.
- [Walk91] WALKER, N.; SMELCER, J. B. & NILSEN, E., Optimizing speed and accuracy of menu selection: A comparison of walking and pull-down menus, *International Journal of Man-Machine Studies*, 35, pp. 871-890.
- [Walt86] WALTON, M., *The Deming Management Method*. A Perigee Book, Putnam, 1986.

- [Wano72] WANOUS, J. P. & LAWLER, E. E., Measurement and Meaning of Job Satisfaction, **Journal of Applied Psychology**, 56(2), April 1972, pp. 95-105.
- [Wass89] WASSERMAN, A. S., Redesigning XEROX: A design strategy based on operability. In: Klemmer, E. T. (Ed.), **Ergonomics: Harness the Power of Human Factors in Your Business**. Ablex, Norwood, NJ, 1989, pp. 7-44.
- [Wast90] WASTELL, D., Mental effort and task performance: Towards a psychophysiology of human-computer interaction. In: IFIP INTERACT'90 Third International Conference on Human-Computer Interaction. **Proceedings**. Cambridge, UK, 27-31 August, 1990, pp. 107-112.
- [Wats97] WATSON, B.; SPAULDING, V.; WALKER, N. & RIBARSKY, W., Evaluation of the Effects of Frame Time Variation on VR Task Performance. In: Virtual Reality Annual International Symposium - VRAIS '97. **Proceedings**. Albuquerque, New Mexico, March 1-5, 1997, pp. 38-44
- [Web00] WEB, **Webopedia Webreference: software engineering**. Dicionário *online* disponível webopedia.internet.com/.
- [Weiz86] WEIZEBAUM, J., ELIZA - a computer program for the study of natural language communications between man and machine, **Communications of the ACM**, 9, September 1986, pp. 36-45.
- [Whar92] WHARTON, C.; BRADFORD, J.; JEFFRIES, R. & FRANZKE, M., Applying Cognitive Walkthroughs to More Complex Interfaces: Experiences, Issues and Recommendations. In: ACM CHI'92 Conference, 1992. **Proceedings**. Monterey, California, May 3-7 1992, pp. 381-388.
- [Whar94] WHARTON, C.; RIEMAN, J.; LEWIS, C. & POLSON, P., The cognitive walkthrough method: A practitioner's guide. In: Jakob Nielsen and Robert L. Mack, **Usability Inspection Methods**. John Wiley & Sons, Inc. New York, 1994, pp. 105-140.
- [Whit91] WHITEFIELD, A.; WILSON, F. & DOWELL, J., A framework for human factors evaluation, **Behaviour & Information Technology**, 10(1), January/February 1991, pp. 65-79.
- [Whit88] WHITESIDE, J., BENNETT, J. & HOLTZBLATT, K., Usability Engineering: Our Experience and Evolution. In: M. Helander (Ed.), **Handbook of Human-Computer Interaction**. Elsevier Science (North Holland), Amsterdam, 1988, pp. 791-817.
- [Wiki94] WIKLUND, M. E., **Usability in practice: How companies develop user friendly products**. Academic Press, Boston, 1994
- [Wild93] WILDEMUTH, B. M., Post-positivist research: two examples of methodological pluralism. **Library Quarterly**, 63, 1993, pp. 450-468.
- [Will97] WILLIAMS, M. G. & BUEHLER, J. N., A study of program entry time predictions for application-specific visual and textual languages. 7th Workshop on Empirical Studies of Programmers. **Proceedings**. Alexandria, VA, October 24 - 26 1997, pp. 209-223
- [Wils98] WILSON, CHAUNCEY, Pros and Cons of Co-participation in Usability Studies, **Usability Interface**, 4(4), April 1998. (Versão *online* disponível em stc.org/pics/usability/newsletter/9804-coparticipation.html)
- [Wino95] WINOGRAD, T., Environments for designing, **Communications of the ACM**, 38(6), June 1995, pp. 65-74.

- [Wood96] WOOD, L., The Ethnographic Interview in User-Centered Work/Task Analysis. In: D. R. Wixon, & J. Ramey, **Field methods Casebook for Software Design**, John Wiley & Sons – New York, 1996, pp. 35-56.
- [Wrig91] WRIGHT, P. C. & MONK, A. F., A cost-effective evaluation method for use by designers, **International Journal of Man-Machine Studies**, 35, 1991, pp. 891-912.
- [Wrig92] WRIGHT, P. C. & CONVERSE, S. A., Method bias and concurrent verbal protocol in software usability testing. In: Human Factors Society 36th Annual Meeting, 1992. **Proceedings**. 1991, pp. 1220-1224.
- [Yama95] YAMADA, S., HONG, J.-K. & SUGITA, S., Development and Evaluation of Hypermedia for Museum Education: Validation of Metrics, **ACM Transactions on Computer-Human Interaction**, 2(4), December 1995, pp. 284-307.
- [Yang90] YANG, Y., Interface usability engineering under practical constraints: A case study in the design of undo support. IFIP INTERACT'90 Third International Conference on Human-Computer Interaction, 1990. **Proceedings**. Cambridge, UK, 27-31 August, 1990, pp. 549-554.
- [Yin94] YIN, R. K., **Case study research: Design and methods** (2nd ed., Applied Social Research Methods Series, Vol. 5). Sage Publications. Thousand Oaks, CA, 1994.
- [Your79] YOURDON, E., **Structured Walkthroughs**. Yourdon Press, Englewood Cliffs, N.J., 1979.
- [Zand96] ZANDEN, B. T. V. & VENCKUS, S. A., An Empirical Study of Constraint Applications Usage in Graphical Applications. In: ACM Symposium on User Interface Software and Technology UIST'96. **Proceedings**. Seattle, WA, November 6 - 8 1996, pp. 137-146.
- [Zhan97] ZHANG, G.; GUAN, Z.; DAI, G. & REN, X., Study and Application of Distributed Multimodal User Interface. In: CSCWID'97. **Proceedings**. Kangawa, Japan. 15 - 17 July 1997, pp. 255-262.
- [Zhan98a] ZHANG, Z.; BASILI, V. & SHNEIDERMAN, B., An empirical study of perspective-based usability inspection. In: Human Factors and Ergonomics Society 42nd Annual Meeting. **Proceedings**. Chicago, 1998, pp. 1346-1350.
- [Zhan98b] ZHANG, Z.; BASILI, V. & SHNEIDERMAN, B., Perspective-based usability inspection. In: Usability Professionals' Association Conference, UPA'98. **Proceedings**. Washington DC., 1998, pp. 281-282.
- [Zhan98c] ZHANG, Q.; WOLF, C. G.; DAIJAVAD, S. & TOUMA, M., Talking to Customers on the Web: A Comparison of Three Voices Alternatives. In: ACM 1998 Conference on Computer Supported Cooperative Work CSCW'98. **Proceedings**. Seattle, WA, November 14 - 18 1998, pp. 109-117.
- [Zhan98d] ZHANG, G.; GUAN, Z.; DAI, G. & REN, X., A comparison of four interaction modes for CAD systems. In: 3rd Asian Pacific Computer & Human Interaction APCHI'98. **Proceedings**. Kangawa, Japan. 15 - 17 July 1998, pp. 82-88.

AVALIAÇÃO DE INTERFACES HOMEM-MÁQUINA

CENTRADA NA INTERAÇÃO HOMEM-MÁQUINA

- Ensaios de Usabilidade
(*Usability Testing*)
 - ☛ Observações (*Observation*)
 - ☛ Questionários (*Questionnaires*)
 - ☛ Entrevistas (*Interviews*)
 - ☛ Verbalização de Procedimentos
(*Thinking Aloud*)
 - ☛ Interação Construtiva
(*Constructive Interaction*)
 - ☛ Ensaio Retrospectivo
(*Retrospective Testing*)
 - ☛ Captura Automática
(*Automatic Logging*)
 - ☛ Discussões em Grupo
(*Focus Groups*)
 - ☛ Retorno das Opiniões do Usuário
(*User Feedback*)
- Avaliação Empírica
(*Empirical Evaluation*)

CENTRADA NA INSPEÇÃO DA DA INTERFACE POR ESPECIALISTAS

- Avaliação Heurística
(*Heuristic Evaluation*)
- Uso de Diretrizes
(*Guidelines Use*)
- Revisões Sistemáticas
(*Walkthroughs*)
- Inspeções de Consistência
(*Consistency Inspections*)
- Inspeções de Padrões
(*Standards Inspections*)
- Inspeções de Características
(*Features Inspections*)
- Inspeções Formais
(*Formal Inspections*)

CENTRADA EM MODELOS

- Cognitivos
(*GOMS [Card83], CCT [Kier85]*)
- Lingüísticos
(*BNF [Rels81], TAG [Payn86]*)
- Físicos
(*KLM [Card90], Tri-estado [Buxt90]*)
- de Interação
(*PIF [Dix98]*)

Híbrida (ou Adaptativa)

ANEXO B

Membros Nacionais da ISO e da IEC (1/4)

País	Entidade de Padronização Nacional	
	Representação na ISO	Representação na IEC
África do Sul	South African Bureau of Standards (SABS)	South African National Committee
Albânia	General Directorate of Standardization (DPS)	
Alemanha	Deutsches Institut für Normung (DIN)	Deutsches Komitee (DKE)
Arábia Saudita	Saudi Arabian Standards Organization (SASO)	Saudi Arabian National Committee
Argélia	Institut Algérien de Normalisation (IANOR)	
Argentina	Instituto Argentino de Normalización (IRAM)	
Armênia	Department for Standardization, Metrology and Certification (SARM)	
Austrália	Standards Australia International Ltd. (SAI)	Australian National Committee
Áustria	Austrian Standards Institute (ON)	Austrian Electrotechnical Committee (ÖVE)
Azerbaijão	Azerbaijan State Standardization and Metrology Centre (AZGOST)	
Bahrain	Directorate of Standards and Metrology (BSMD)	
Bangladesh	Bangladesh Standards and Testing Institution (BSTI)	
Barbados	Barbados National Standards Institution (BNSI)	
Bielorrússia	Committee for Standardization, Metrology and Certification (BELST)	Belarus National Committee
Bélgica	The Belgian Institution for Standardization (IBN)	Comité Electrotechnique Belge (CEB)
Benin	Direction de la Promotion de la Qualité et du Conditionnement des Produits (DPQC)	
Bolívia	Instituto Boliviano de Normalización y Calidad (IBNORMA)	
Bósnia e Herzegovina	Institute for Standardization, Metrology and Patents of Bosnia and Herzegovina (BASMP)	National Committee of Bosnia and Herzegovina
Botswana	Botswana Bureau of Standards (BOBS)	
Brasil	Associação Brasileira de Normas Técnicas (ABNT)	Brazilian National Committee (Suspensão em 01/10/1999)
Brunel	Construction Planning and Research Unit (CPRU)	
Bulgária	State Agency for Standardization and Metrology (BDS)	Bulgarian National Committee
Burkina Faso	Direction de la Normalisation et de la Promotion de la Qualité (FASONORM)	
Camboja	Industrial Standards Bureau (ISB)	
Canadá	Standards Council of Canada (SCC)	Canadian National Committee
Cazaquistão	Committee for Standardization, Metrology and Certification (KAZMEMST)	
Chile	Instituto Nacional de Normalización (INN)	
China	China Standards Information Center (CSSN) China State Bureau of Quality and Technical Supervision (CSBTS)	Chinese National Committee
Chipre	Cyprus Organization for Standards and Control of Quality (CYS)	National Committee of Cyprus
Colômbia	Instituto Colombiano de Normas Técnicas y Certificación (ICONTEC)	
Comores	Service de la Normalisation / Qualité (CSNQ)	
Costa do Marfim	Organisme Ivoirien de Normalisation et de Certification (CODINORM)	
Costa Rica	Instituto de Normas Técnicas de Costa Rica (INTECO)	(IEC Pre-Associate Member, Admissão em 02/2000)
Croácia	State Office for Standardization and Metrology (DZNM)	
Cuba	Oficina Nacional de Normalización (NC)	(IEC Pre-Associate Member, Admissão em 01/1998)
Dinamarca	Dansk Standard (DS)	
Egito	Egyptian Organization for Standardization and Quality Control (EOS)	The Egyptian National Committee
El Salvador	Consejo Nacional de Ciencia y Tecnología (CONACYT)	
Emirados Árabes	Directorate of Standardization and Metrology (SSUAE)	
Equador	Instituto Ecuatoriano de Normalización (INEN)	
Eslováquia	Slovak Office of Standards, Metrology and Testing (SUTN)	Slovensky Elektrotechnický Výbor (SEV)

ANEXO B

Membros Nacionais da ISO e da IEC (2/4)

País	Entidade de Padronização Nacional	
	Representação na ISO	Representação na IEC
Eslovênia	Standards and Metrology Institute (SMIS)	Slovenian IEC National Committee
Espanha	Asociación Española de Normalización y Certificación (AENOR)	Comité Nacional Español de la CEI
Estados Unidos	American National Standards Institute (ANSI)	U.S. National Committee
Estônia	Eesti Standardikeskus (ESK)	Estonian National Committee
Etiópia	Quality and Standards Authority of Ethiopia (QSAE)	
Federação Russa	State Committee of the Russian Federation for Standardization, Metrology and Certification (GOST-R)	Russian Federation Committee
Fiji	Fiji Trade Standards and Quality Control Office (FTSQCO)	
Filipinas	Bureau of Product Standards (BPS)	
Finlândia	Finnish Standards Association (SFS)	Finnish Electrotechnical Standards Association (SESKO)
França	Association Française de Normalisation (AFNOR)	Union Technique de l'Électricité (UTE)
Gana	Ghana Standards Board (GSB)	
Granada	Grenada Bureau of Standards (GDBS)	
Grécia	Hellenic Organization for Standardization (ELOT)	
Guatemala	Comisión Guatemalteca de Normas (COGUANOR)	
Guiana	Guyana National Bureau of Standard (GNBS)	
Guiné	Institut National de Normalisation et de Métrologie (INNEM)	
Holanda	Nederlands Normalisatie-Instituut (NEN)	Netherlands National Committee (NEC)
Honduras	Consejo Hondureño de Ciencia y Tecnología (COHCIT)	
Hong Kong, China	Industry Department (ITCHK SAR)	
Hungria	Magyar Szabványügyi Testület (MSZT)	
Índia	Bureau of Indian Standards (BIS)	
Indonésia	Badan Standardisasi Nasional (BSN)	
Irlanda	National Standards Authority of Ireland (NSAI)	Electro-Technical Council of Ireland Ltd.
Islândia	Icelandic Council for Standardization (STRI)	
Israel	The Standards Institution of Israel (SII)	
Itália	Ente Nazionale Italiano di Unificazione (UNI)	Comitato Elettrotecnico Italiano (CEI)
Iugoslávia	Savezni Zavod za Standardizaciju (SZS)	Federal Institution for Standardization
Jamaica	Jamaica Bureau of Standards (JBS)	
Japão	Japan Industrial Standards Committee (JISC)	
Jordânia	Jordan Institution for Standards and Metrology (JISM)	
Kuwait	Standards and Industrial Services Affairs (KOWSMD)	
Látvia	Latvian Standards (LVS)	Latvian National Committee
Lesotho	Lesotho Standards and Quality Assurance Section (LSQAS)	
Líbano	Lebanese Standards Institution (LIBNOR)	
Líbia	Libyan National Centre for Standardization and Metrology (LNCMS)	
Lituânia	Lithuanian Standards Board (LST)	
Luxemburgo	Service de l'Energie de l'Etat, Département Normalisation (SEE)	Comité National du Luxembourg
Macedônia	Zavod za Standardizacija i Metrologija (ZSM)	
Madagáscar	Bureau de Normes de Madagascar (BNM)	
Malásia	Department of Standards Malaysia (DSM)	Malaysian National Committee
Malawi	Malawi Bureau of Standards (MBS)	
Mali	Direction Nationale des Industries (MLIDNI)	
Malta	Malta Standardisation Authority (MAS)	Maltese National Committee
Marrocos	Service de Normalisation Industrielle Marocaine (SNIMA)	
Maurício	Mauritius Standards Bureau (MSB)	

ANEXO B

Membros Nacionais da ISO e da IEC (3/4)







País	Entidade de Padronização Nacional	
	Representação na ISO	Representação na IEC
México	Dirección General de Normas (DGN)	
Moçambique	Instituto Nacional de Normalização e Qualidade (INNOQ)	
Moldávia	Department of Standards, Metrology and Technical Supervision (MOLDST)	
Mongólia	Mongolian National Centre for Standardization and Metrology (MNCMS)	
Namíbia	Namibia Standards Information and Quality Office (NSIQO)	
Nepal	Nepal Bureau of Standards and Metrology (NBSM)	
Nicarágua	Dirección General de Ciencia y Tecnología (DGCYT)	
Noruega	Norges Standardiseringsforbund (NSF)	Norsk Electroteknisk Komite (NEK)
Nova Zelândia	Standards New Zealand (SNZ)	New Zealand Electrotechnical Committee
Oman	Directorate General for Specifications and Measurements (DGSM)	
Panamá	Comisión Panameña de Normas Industriales y Técnicas (COPANIT)	
Papuásia	National Institute of Standards and Industrial Technology (NISIT)	
Paquistão	Pakistan Standards Institution (PSI)	Pakistan National Committee
Paraguai	Instituto Nacional de Tecnología y Normalización (INTN)	
Peru	Instituto Nacional de Defensa de la Competencia y de la Protección de la Propiedad Intelectual (INDECOPI)	
Poónia	Polish Committee for Standardization (PKN)	Polish National Committee
Portugal	Instituto Português da Qualidade (IPQ)	Portuguese National Committee
Qatar	Dpt. of Standards, Measurements and Consumer Protection (QS)	
Quênia	Kenya Bureau of Standards (KEBS)	(IEC Pre-Associate Member, Admissão em 02/1997)
Quirguistão	State Inspection for Standardization and Metrology (KYRGYZST)	
Reino Unido	British Standards Institution (BSI)	British Electrotechnical Committee
Rep. da Coreia	Korean Agency for Technology and Standards (KATS)	Korean National Committee
Rep. Dem. do Congo	Office Congolais de Contrôle (OCC)	
Rep. Dem. Pop. da Coreia	Committee for Standardization of the Democratic People's Republic of Korea (CSK)	
Rep. Dominicana	Dirección General de Normas y Sistemas de Calidad (DIGENOR)	
Rep. dos Camarões	Cellule de la Normalisation et de la Qualité (CCNQ)	
Rep. Islâmica do Irã	Institute of Standards and Industrial Research of Iran (IRISI)	Iranian National Committee
Rep. Tcheca	Czech Standards Institute (CSNI)	
Rep. Unida da Tanzânia	Tanzania Bureau of Standards (TBS)	
Romênia	Asociația de Standardizare din România (ASRO)	Romanian National Committee (CER)
Santa Lucía	Saint Lucia Bureau of Standards (SLBS)	
Seychelles	Seychelles Bureau of Standards (SBS)	
Singapura	Singapore Productivity and Standards Board (PSB)	Singapore National Committee
Síria	Syrian Arab Organization for Standardization and Metrology (SASMO)	
Sri Lanka	Sri Lanka Standards Institution (SLSI)	
Sudão	Sudanese Standards and Metrology Organization (SSMO)	
Suécia	Standardiseringsen i Sverige (SIS)	Svenska Elektriska Kommissionen (SEK)
Sulça	Swiss Association for Standardization (SNV)	Swiss Electrotechnical Committee (CES)
Tailândia	Thai Industrial Standards Institute (TISI)	Thai National Committee
Trinidade e Tobago	Trinidad and Tobago Bureau of Standards (TTBS)	
Tunísia	Institut National de la Normalisation et de la Propriété Industrielle (INNORPI)	Comité National Tunisien
Turquia	Türk Standardları Enstitüsü (TSE)	Turkish National Committee
Turquemenistão	Major State Inspection "Turkmenstandartary" (MSIT)	

ANEXO B



Membros Nacionais da ISO e da IEC (4/4)

País	Entidade de Padronização Nacional	
	Representação na ISO	Representação na IEC
Ucrânia	State Committee of Ukraine for Standardization, Metrology and Certification (DSTU)	Ukrainian National Committee
Uganda	Uganda National Bureau of Standards (UNBS)	
Urugual	Instituto Uruguayo de Normas Técnicas (UNIT)	(IEC Pre-Associate Member, Admissão em 09/1996)
Usbequistão	Uzbek State Centre for Standardization, Metrology and Certification (USGOST)	
Venezuela	Fondo para la Normalización y Certificación de la Calidad (FONDONORMA)	
Vietnã	Directorate for Standards and Quality (TCVN)	
Zimbabwe	Standards Association of Zimbabwe (SAZ)	

LEGENDA DE CATEGORIAS DE MEMBROS ISO E IEC

 ISO Full Member	Corpo nacional de padronização mais representativo. Participação ativa em quaisquer dos comitês técnicos e de Inspeção da ISO. Plenos direitos em votações.
 ISO Correspondent Member	Organização nacional com atividades de padronização em nível nacional ainda não plenamente desenvolvidas. Não participação em comitês técnicos e de Inspeção da ISO. Direito de recebimento de todas as informações sobre atividades de padronização de seu interesse.
 ISO Subscriber Member	País detentor de economias em escalas bastante reduzidas. Taxas de adesão reduzidas. Contato com atividades de padronização em nível Internacional.
 IEC Full Member	Participação ativa e total em atividades de padronização Internacional através de comitê nacional. Plenos direitos em votações.
 IEC Associate Member	Participação limitada em trabalhos de padronização. País com recursos limitados. Status de observador. Participação em todas os encontros. Sem direito de voto.
 IEC Pre-Associate Member	País com suporte para a formação de um comitê nacional em eletrotécnica do IEC Central Office ou do comitê nacional de qualquer país membro vizinho, com o objetivo de tornar-se um Associate Member após um período de cinco anos.

LEGENDA COMPLEMENTAR

 Non Associate	País não sócio da IEC.
 Suspended Full Member	Não se trata de uma categoria, mas de uma <u>condição</u> . Segundo a IEC [IEC00], a comissão nacional brasileira (Brazilian National Committee) se encontra suspensa desde 01/10/1999.

ANEXO C

Padrões ISO para a Ergonomia da Interação Homem-Máquina (1/3)

Padrões ISO para a Ergonomia da Interação Homem-Máquina					
Referência	Especificação	Comitê Técnico/ Sub-comitê Responsável		Status	Data de atribuição do Status
		TC	SC		
ISO/AWI 20282	Ergonomic guiding principles - Classification of usability of man/machine interfaces - Evaluation methods (Ed. 1)	159	1	20.00	14/04/2000
ISO/AWI TS 19358	Ergonomic assessment of speech technology systems (automatic speech recognition and synthesis) (Ed. 1)	159	5	20.00	30/11/1999
ISO/AWI 18789	Ergonomics of human system interaction – Ergonomic requirements and measurement techniques for electronic visual displays (Ed. 1)	159	4	20.00	21/12/1999
ISO/TR 18529:2000	Ergonomics – Ergonomics of human-system interaction – Human-centred lifecycle process descriptions (available in English only) (Ed. 1, 28 p, N)	159	4	60.60	22/06/2000
ISO/WD TS 16982	Ergonomics of human-system interaction – Usability methods supporting human centered design (Ed. 1)	159	4	20.20	02/06/2000
ISO/WD TR 16071	Ergonomics of human-system interaction – Guidance on accessibility for human-computer interfaces (Ed. 1)	159	4	20.20	02/05/2000
ISO/IEC AWD 16005	Keyboard layouts for text and office systems – Transportable computers (Ed. 1)	JTC1	-	20.00	03/10/1997
ISO/IEC CD 15441	Keyboard layouts for text and office systems – Transportable computers (Ed. 1)	JTC1	-	30.30	02/06/1997
ISO/IEC 15412:1999	Information technology – Portable computer keyboard layouts (available in English only) (Ed. 1, 4 p, B)	JTC1	35	60.60	09/09/1999
ISO/IEC 15411:1999	Information technology – Segmented keyboard layouts (available in English only) (Ed. 1, 15 p, H)	JTC1	35	60.60	14/10/1999
ISO 15187	Manipulating industrial robots – Graphical user interfaces for programming and operation of robots (GUI-R) (Ed. 1, 22 p)	184	2	60.00	25/09/2000
ISO/IEC 14598-1:1999	Information technology – Software product evaluation – Part 1: General overview (available in English only) (Ed. 1, 19 p, K)	JTC1	7	60.60	22/04/1999
ISO/IEC 14598-2:2000	Software engineering – Product evaluation – Part 2: Planning and management (available in English only) (Ed. 1, 12 p, F)	JTC1	7	60.60	10/02/2000
ISO/IEC 14598-3:2000	Software engineering – Product evaluation – Part 3: Process for developers (available in English only) (Ed. 1, 16 p, H)	JTC1	7	60.60	10/02/2000
ISO/IEC 14598-3:2000/ CD Amd 1	Software engineering – Product evaluation – Part 3: Process for developers (available in English only) (Ed. 1))	JTC1	7	30.99	18/05/1999
ISO/IEC 14598-4:1999	Software engineering – Product evaluation – Part 4: Process for acquirers (available in English only) (Ed. 1, 34 p, Q)	JTC1	7	60.60	14/10/1999
ISO/IEC 14598-5:1998	Information technology – Software product evaluation – Part 5: Process for evaluators (available in English only) (Ed. 1, 35 p, Q)	JTC1	7	60.60	02/07/1998
ISO/IEC 14598-6	Information technology – Software product evaluation – Part 5: Process for evaluators (available in English only) (Ed. 1, 35 p, Q)	JTC1	7	60.00	29/11/1999
ISO/DIS 14915-1	Software ergonomics for multimedia user interfaces – Part 1: Design principles and framework (Ed. 1, 13 p, *)	159	4	40.60	18/07/2000

Padrões ISO para a Ergonomia da Interação Homem-Máquina (2/3)

Referência	Especificação	Comitê Técnico/ Sub-comitê Responsável		Status	Data de atribuição do Status
		TC	SC		
ISO/CD 14915-2	Software ergonomics for multimedia user interfaces – Part 2: Multimedia control and navigation (<i>Ed. 1</i>)	159	4	30.20	02/06/2000
ISO/DIS 14915-3	Software ergonomics for multimedia user interfaces – Part 3: Media selection and combination (<i>Ed. 1, 45 p, *</i>)	159	4	40.60	18/07/2000
ISO/IEC 14754:1999	Information technology – Pen-Based Interfaces – Common gestures for Text Editing with Pen-Based Systems (available in English only) (<i>Ed. 1, 15 p, H</i>)	JTC1	35	60.60	27/05/1999
ISO/IEC 13714:1995	Information technology – Document processing and related communication – User interface to telephone-based services – Voice messaging applications (available in English only) (<i>Ed. 1, 60 p, U</i>)	JTC1	-	90.20	08/06/2000
ISO/FDIS 13406-1	Ergonomic requirements for work with visual displays based on flat panels – Part 1: Introduction (<i>Ed. 1, 3 p, B</i>)	159	4	50.20	31/08/2000
ISO/FDIS 13406-2	Ergonomic requirements for work with visual displays based on flat panels – Part 2: Ergonomic requirements for flat panel displays (<i>Ed. 1, 146 p</i>)	159	4	60.60	07/10/1999
ISO 13407:1999	Human-centred design processes for interactive systems (<i>Ed. 1, 26 p, M</i>)	159	4	60.60	10/06/1999
ISO/IEC 11581-1:2000	Information technology – User system interfaces and symbols – Icon symbols and functions – Part 1: Icons – General (available in English only) (<i>Ed. 1, 9 p, E</i>)	JTC1	35	60.60	23/03/2000
ISO/IEC 11581-2:2000	Information technology – User system interfaces and symbols – Icon symbols and functions – Part 2: Object icons (available in English only) (<i>Ed. 1, 25 p, M</i>)	JTC1	35	60.60	23/03/2000
ISO/IEC 11581-3:2000	Information technology – User system interfaces and symbols – Icon symbols and functions – Part 3: Pointer icons (available in English only) (<i>Ed. 1, 7 p, D</i>)	JTC1	35	60.60	30/03/2000
ISO/IEC CD 11581-4	Information technology – User system interfaces and symbols – Icon symbols and functions – Part 4: Control icons (<i>Ed. 1</i>)	JTC1	35	30.20	22/06/1999
ISO/IEC FCD 11581-5	Information technology – User system interfaces and symbols – Icon symbols and functions – Part 5: Tool icons (<i>Ed. 1</i>)	JTC1	35	40.20	09/08/2000
ISO/IEC 11581-6:1999	Information technology – User system interfaces and symbols – Icon symbols and functions – Part 6: Action icons (available in English only) (<i>Ed. 1, 15 p, H</i>)	JTC1	35	60.60	21/01/1999
ISO/IEC AW11588-9	Information technology – Message Handling Systems (MHS) management – Part 9: User agent entity (<i>Ed. 1</i>)	JTC1	6	20.00	01/01/1994
ISO/WD 11064-5	Ergonomic design of control centres – Part 5: Human-system interfaces (<i>Ed. 1</i>)	159	4	20.20	01/03/1999
ISO/IEC 10741-1:1995	Information technology – User system interfaces – Dialogue interaction – Part 1: Cursor control for text editing (available in English only) (<i>Ed. 1, 9 p, E</i>)	JTC1	35	90.20	08/06/2000

ANEXO C

Padrões ISO para a Ergonomia da Interação Homem-Máquina (3/3)

Referência	Especificação	Comitê Técnico/ Sub-comitê Responsável		Status	Data de atribuição do Status
		TC	SC		
ISO/IEC AWI 10741-2	Information technology – User system interfaces - Dialogue interaction – Part 2: Forms entry and spreadsheets (Ed. 1)	JTC1	35	20.00	01/11/1993
ISO/IEC 9995-1:1994	Information technology – Keyboard layouts for text and office systems – Part 1: General principles governing keyboard layouts (available in English only) (Ed. 1, 11 p, F)	JTC1	35	90.60	30/10/1999
ISO/IEC 9995-2:1994	Information technology – Keyboard layouts for text and office systems – Part 2: Alphanumeric section (available in English only) (Ed. 1, 7 p, D)	JTC1	35	90.60	30/10/1999
ISO/IEC 9995-3:1994	Information technology – Keyboard layouts for text and office systems – Part 3: Complementary layouts of the alphanumeric zone of the alphanumeric section (Ed. 1, 5 p, C)	JTC1	35	90.60	30/10/1999
ISO/IEC 9995-4:1994	Information technology – Keyboard layouts for text and office systems – Part 4: Numeric section (available in English only) (Ed. 1, 7 p, D)	JTC1	35	90.60	30/10/1999
ISO/IEC 9995-5:1994	Information technology – Keyboard layouts for text and office systems – Part 5: Editing section (available in English only) (Ed. 1, 7 p, D)	JTC1	35	90.60	30/10/1999
ISO/IEC 9995-6:1994	Information technology – Keyboard layouts for text and office systems – Part 6: Function section (available in English only) (Ed. 1, 2 p, A)	JTC1	35	90.60	30/10/1999
ISO/IEC 9995-7:1994	Information technology – Keyboard layouts for text and office systems – Part 7: Symbols used to represent functions (Ed. 1, 14 p, G)	JTC1	35	90.60	30/10/1999
ISO/IEC 9995-8:1994	Information technology – Keyboard layouts for text and office systems – Part 8: Allocation of letters to the keys of a numeric keypad (available in English only) (Ed. 1, 3 p, B)	JTC1	35	90.60	30/10/1999
ISO 9355-1:1999	Ergonomic requirements for the design of displays and control actuators – Part 1: Human interactions with displays and control actuators (Ed. 1, 14 p, G)	159	4	60.60	02/12/1999
ISO 9355-2:1999	Ergonomic requirements for the design of displays and control actuators – Part 2: Displays (Ed. 1, 18 p, J)	159	4	60.60	02/12/1999
ISO 9127:1988	Information processing systems – User documentation and cover information for consumer software packages (Ed. 1, 7 p, D)	JTC1	7	90.60	30/10/1999
ISO 6385:1981	Ergonomic principles in the design of work systems (Ed. 1, 4 p, B)	159	-	90.92	01/06/1991
ISO 2382-21:1985	Data processing – Vocabulary – Part 21: Interfaces between process computer systems and technical processes (Ed. 1, 16 p, H)	JTC1	-	90.20	08/06/2000

ANEXO D

Produto Avaliado: **MATLAB v. 5.3.0 (The MathWorks Inc.)**

Desenvolvido por *The Math Works Inc.*, o **MATLAB** é um sistema que integra computação, visualização e programação em um ambiente no qual problemas e soluções são expressos em notação matemática familiar. Usos típicos incluem o desenvolvimento de algoritmos; a modelagem, simulação e prototipagem de sistemas; a análise, exploração e visualização de dados; e o desenvolvimento de aplicações, incluindo implementação de interfaces gráficas com o usuário.

O ambiente **MATLAB** contém um conjunto de variáveis (denominadas *arrays*), o qual não requer dimensionamento e pode ser manipulado a partir da linha de comandos. Isto possibilita a resolução de diversos problemas técnicos de computação, especialmente aqueles associados formulações matriciais e vetoriais [Hans97]. O ambiente de desenvolvimento do **MATLAB 5.3** inclui, dentre diversos recursos:

- Suporte para tipos inteiros de dados;
- Operações envolvendo matrizes esparsas;
- Análise numérica de dados;
- Ferramentas de development de ambientes;
- Documentação *online*;
- Janela de processamento e visualização de representações gráficas;
- Editor de plotagem de gráficos;
- Funções de visualização 3D e de plotagem de representações gráficas;
- Suporte para imagens PNG (*Portable Network Graphics*); e
- Modos de interação com o usuário via menus, linguagem de comandos e manipulação direta.

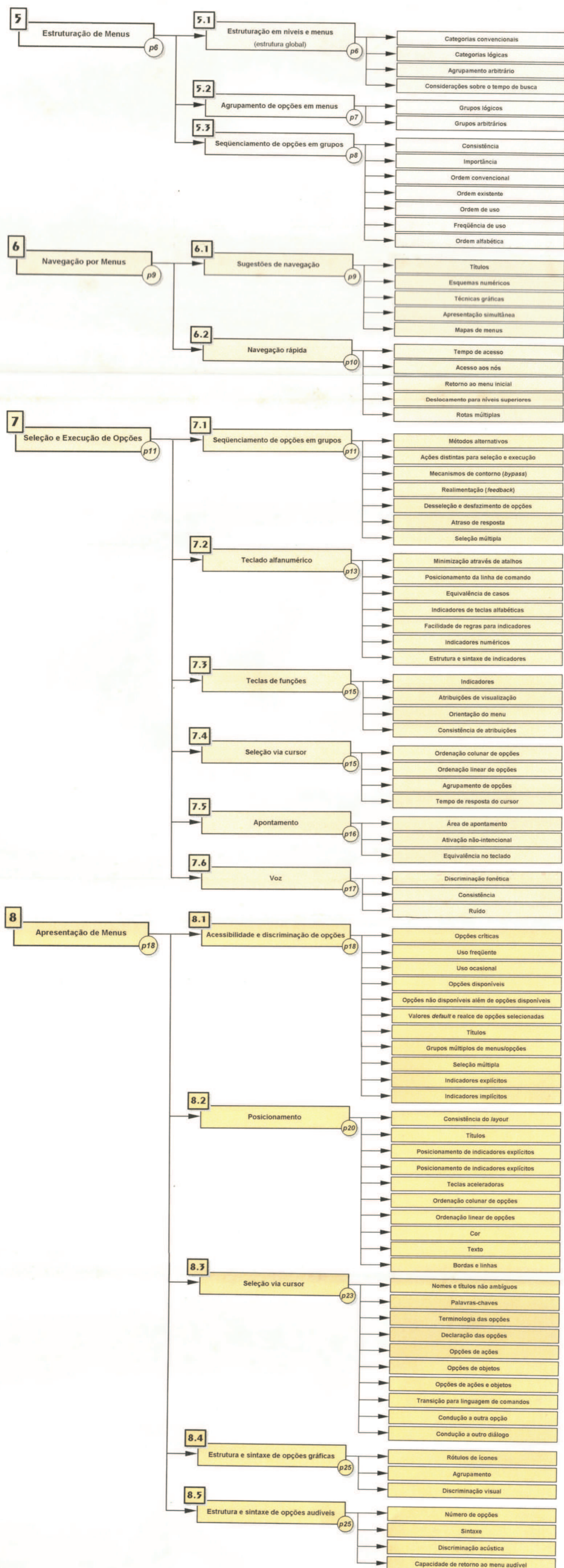
Além de suas funcionalidades, o sistema integra uma série de caixas de ferramentas (*toolboxes*) que suportam atividades em diversas áreas, e.g redes neurais, lógica *fuzzy*, wavelets, estatística, processamento de sinais e imagens e comunicações. São as seguintes as caixas de ferramentas desenvolvidas para o **MATLAB**:

- **Comunicações** (*Communications*)
- **Sistemas de Controle** (*Control System*)

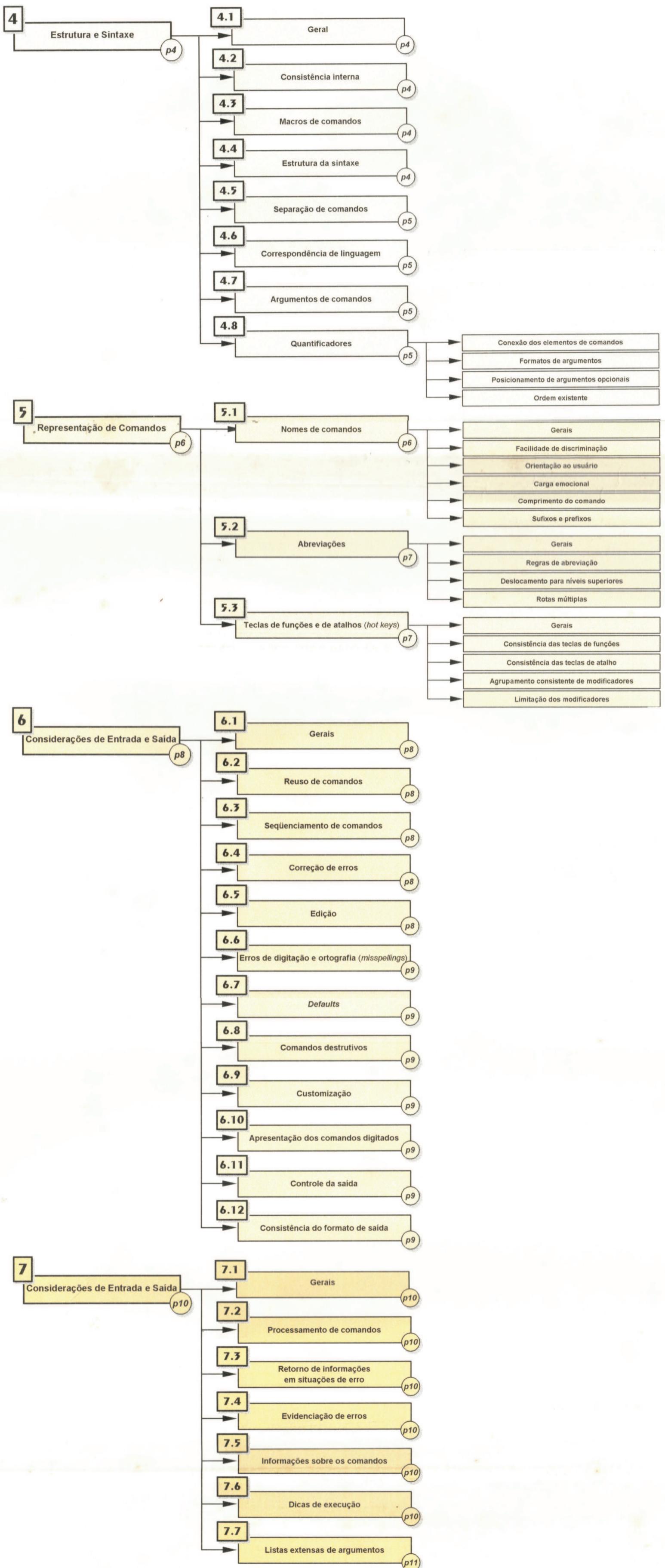
- Identificação de Sistemas no Domínio da Frequência (*Frequency Domain System Identification*)
- Lógica Fuzzy (*Fuzzy Logic*)
- Análise Espectral de Ordem Superior (*Higher-Order Spectral Analysis*)
- Processamento de Imagens (*Image Processing*)
- Controle LMI (*LMI Control*)
- Mapeamento (*Mapping*)
- Controle de Modelos Preditivos (*Model Predictive Control*)
- μ -Análise e Síntese (*μ -Analysis and Synthesis*)
- Fundamentos de NAG (NAG Foundation)
- Redes Neurais (*Neural Network*)
- Otimização (*Optimization*)
- Equações Diferenciais Parciais (*Partial Differential Equation*)
- Projeto de Controle QFT (*QFT Control Design*)
- Controle Robusto (*Robust Control*)
- Processamento de Sinais (*Signal Processing*)
- *Splines*
- Estatística (*Statistics*)
- Matemática Simbólica e Matemática Simbólica Estendida (*Symbolic Math and Extended Symbolic Math*)
- System Identification (*System Identification*)
- *Wavelet*

O *MATLAB* está associado ao *Simulink*, um sistema interativo destinado à simulação de sistemas dinâmicos não-lineares. O *Simulink* possibilita a modelagem de sistemas a partir da construção de diagramas de blocos, diretamente na tela, e da manipulação dinâmica dos blocos. Este sistema suporta a modelagem de sistemas lineares, não-lineares, discretos ou contínuos no tempo e multivariáveis.

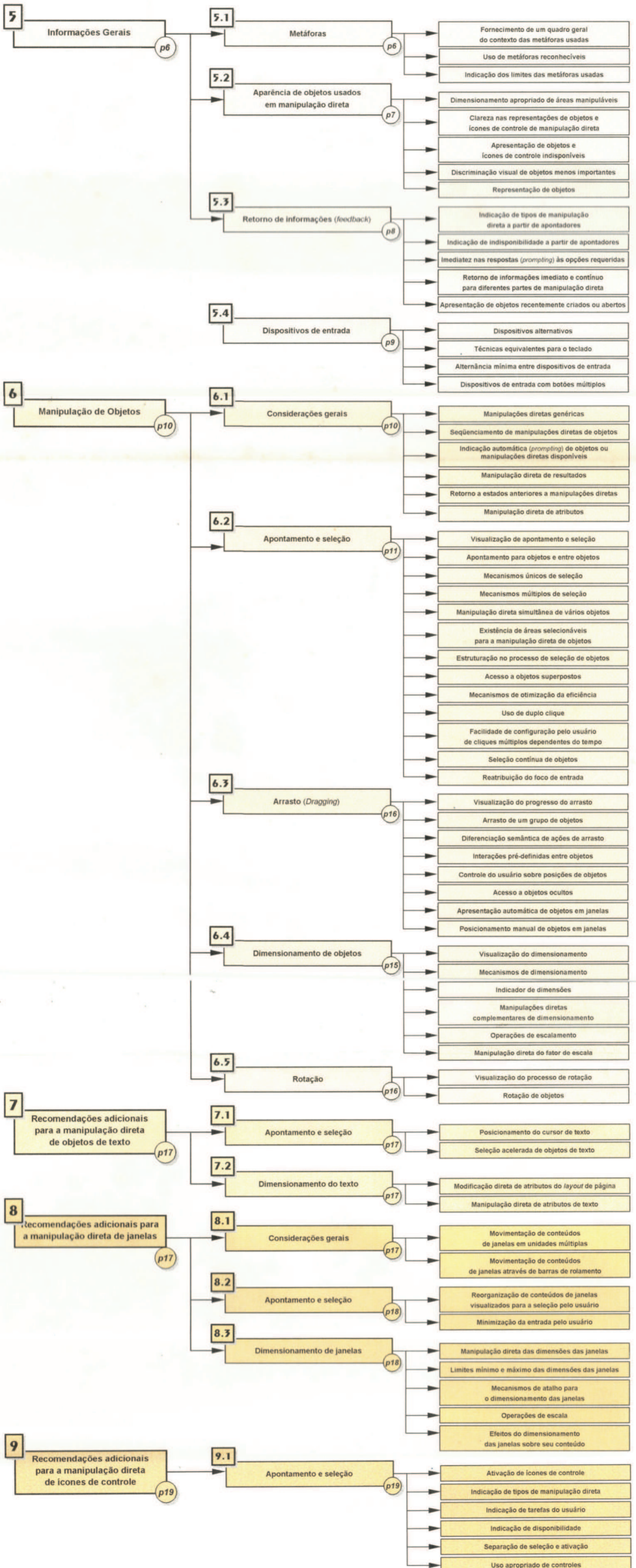
E1 - Diagrama de Seções e Sub-seções do ISO 9241-14



E2 - Diagrama de Seções e Sub-seções do ISO 9241-15



E3 - Diagrama de Seções e Sub-seções do ISO 9241-16



Recomendações	Aplicabilidade								Adoção						Comentários (incluindo fontes)		
	Resultados		Método usado						Método usado							Resultados	
	S	N	A	E	O	AA	AE	MD	M	O	E	AA	AE	MD		P	F
8.2.3 Designadores explícitos Posicionados à esquerda do nome da opção (separados deste por 2 ou 3 caracteres de espaçamento) e																	
8.2.4 Teclas aceleradoras Códigos posicionados à direita do nome da opção (e preferencialmente alinhados pela direita) e																	
Opções em colunas																	
a) Espaçamento – se disponíveis, as opções deverão ser apresentadas em espaço vertical duplo ou																	
b) Espaçamento simples – se as opções forem apresentadas em espaço simples, usar letras minúsculas ou minúsculas com a letra inicial maiúscula e																	
8.2.5																	
c) Grupos de opções – separados por um espaçamento vertical que seja uma vez e meia a duas vezes igual ao espaçamento vertical das opções existentes em cada grupo e																	
d) Alinhamento – As opções deverão ser alinhadas pela esquerda (<i>flush left</i>) e																	
e) Múltiplas colunas of opções - deverão ser separadas por, pelo menos, 3 caracteres de espaçamento e																	
f) Designadores sequenciados – designadores numéricos ou alfabéticos deverão ser alinhados sequencialmente em colunas e																	
Opções em linhas Se posicionadas horizontalmente, deverão ser suficientemente separadas para permitir discriminação visual e/ou																	
8.2.6																	
Cor O mesmo código de cores deverá ser adotado para as opções de um grupo particular de opções (limitando-se a 4 cores) e/ou																	
8.2.7																	
Letras Se forem usados estilos e tamanhos de letras, considerar:																	
a) Legibilidade – estilos e tamanhos de letras deverão ser legíveis e distinguíveis e																	
b) Quantidade – combinações únicas de estilos e tamanhos de letras em um menu não deverão exceder três (não contando maiúsculas/minúsculas) e/ou																	
8.2.8																	
8.2.9 Bordas e linhas:																	
a) As bordas e linhas deverão ser mantidas simples e																	
b) As bordas e linhas deverão ser suficientemente separadas das opções de modo a não interferirem com a legibilidade																	
8.3 Estrutura e sintaxe de opções textuais																	
Nomes e títulos não ambíguos Nomes de opções e títulos de grupos de opções deverão ser semanticamente distintos e																	
8.3.1																	
Palavras-chaves:																	
a) Início com palavras-chaves (a menos que seja anti-naturais para o idioma) e																	
b) Palavras-chaves sugestivas deverão ser usadas, enquanto palavras-chaves inócuas evitadas e																	
8.3.2																	
Terminologia das opções Familiar aos usuários e																	
8.3.3																	
Fraseado das opções Conciso e consistente e																	
8.3.4																	
Opções de ações Expressas através de verbos e/ou																	
8.3.5																	
Opções de objetos Expressas através de substantivos (nomes) e/ou																	
8.3.6																	
Opções de ações e objetos Representadas por sintaxe verbo-nominal e																	
8.3.7																	
Transição para linguagem de comandos Uso de maiúsculas e sintaxe dos nomes das opções devem ser consistentes com a linguagem de comandos e																	
8.3.8																	
Condução a outra opção Se uma opção conduz a outro menu ao invés de conduzir à execução, fornecer indicações apropriadas ou																	
8.3.9																	
Condução a outro diálogo Se a opção conduz a outro diálogo, fornecer indicações consistentes.																	
8.3.10																	
8.4 Estrutura e sintaxe de opções gráficas																	
8.4.1																	
Rótulos de ícones Caso seja possível ambiguidade de ícones e																	
8.4.2																	
Agrupamento Ícones de objetos e de ações posicionados em diferentes grupos de um menu e																	
8.4.3																	
Discriminação visual Ícones devem ser selecionados para representar opções visualmente distintas e seu significado deverá ser facilmente reconhecido.																	
8.5 Estrutura e sintaxe de opções audíveis																	
8.5.1																	
Número de opções O menor possível (3 ou 4) e																	
8.5.2																	
Sintaxe A sintaxe de opções/designadores preferida e																	
8.5.3																	
Discriminação acústica Opções do menu de voz compreendidas de itens distintos em termos auditivos e traduzidos por palavras únicas suficientemente espaçadas (no tempo), a fim de permitir a discriminação pelo usuário. e																	
8.5.4																	
Capacidade de repetição Deverá ser oferecida.																	

NOTA Usuários do ISO 9241-14 poderão reproduzir livremente esta lista de verificação, a fim de que ela possa ser utilizada para seu propósito primordial, podendo também publicar posteriormente a lista de verificação preenchida.

Legenda:			
S = Sim (se aplicável)	A = Análise da Documentação do Sistema	AA = Avaliação Analítica	M = Mensuração
N = Não (se não aplicável)	E = Evidência documentada	AE = Avaliação Empírica	P = Passou (atendeu à recomendação)
	O = Observação	DM = Método Diferente	F = Falhou (não atendeu à recomendação)

F3 - Lista de Inspeção de Conformidade ao ISO 9241-16

(2/2)

Recomendações	Aplicabilidade									Adoção						Comentários (incluindo fontes)
	Resultados		Método usado							Método usado				Resultados		
	S	N	A	E	O	AA	AE	MD	M	O	E	AA	AE	MD	P	
8.2 Apontamento e seleção																
8.2.1 Rearranjo do conteúdo visualizado de uma janela de acordo com a seleção do usuário																
8.2.2 Minimização da entrada de dados pelo usuário																
8.3 Dimensionamento de janelas																
8.3.1 Manipulação direta das dimensões de janelas																
8.3.2 Limites mínimo e máximo das dimensões de janelas																
8.3.3 Atalhos para o dimensionamento																
8.3.4 Manipulação de escala																
8.3.5 Efeitos do dimensionamento de uma janela sobre seu conteúdo																
9 Recomendações adicionais para a manipulação direta de ícones de controle																
9.1 Apontamento e seleção																
9.1.1 Ativação de ícones de controle																
9.1.2 Indicação de tipos de manipulação direta																
9.1.3 Indicação de tarefas do usuário																
9.1.4 Indicação de disponibilidade																
9.1.5 Separação entre seleção e ativação																
9.1.6 Uso apropriado de controles																

NOTA Usuários do ISO 9241-16 poderão reproduzir livremente esta lista de verificação, a fim de que ela possa ser utilizada para seu propósito primordial, podendo também publicar posteriormente a lista de verificação preenchida.

Legenda:

S = Sim (se aplicável)

N = Não (se não aplicável)

A = Análise da Documentação do Sistema

E = Evidência documentada

O = Observação

AA = Avaliação Analítica

AE = Avaliação Empírica

DM = Método Diferente

M = Mensuração

P = Passou (atendeu à recomendação)

F = Falhou (não atendeu à recomendação)



DePerUSI[®]



Grupo de Interfaces Homem-Máquina

DELINEAMENTO DO PERFIL DO USUÁRIO DE SISTEMAS INTERATIVOS

Este questionário tem o propósito de coletar informações que possibilitem delinear seu perfil de usuário de sistemas computacionais e, em particular, do PRODUTO sob condições de teste. Por favor analise cada aspecto considerado, selecionando a opção que mais adequa à sua condição de usuário e fornecendo as demais informações solicitadas, quando se fizer necessário.

ASPECTO	OPÇÕES
1. Você é:	<input type="radio"/> NÍVEL ACADÊMICO <input type="radio"/> PÓS-GRADUADO <input type="radio"/> GRADUADO <input type="radio"/> ESTUDANTE DE PÓS-GRADUAÇÃO <input type="radio"/> ESTUDANTE DE GRADUAÇÃO
2. Você é do sexo:	<input type="radio"/> SEXO <input type="radio"/> MASCULINO <input type="radio"/> FEMININO
3. Você é:	<input type="radio"/> DESTREZA MANUAL <input type="radio"/> DESTRO <input type="radio"/> CANHOTO <input type="radio"/> AMBIDESTRO
4. Você usa óculos ou lentes de contato?	<input type="radio"/> USO DE CORRETIVOS VISUAIS <input type="radio"/> SIM <input type="radio"/> NÃO
5. Você pertence à faixa etária de:	<input type="radio"/> FAIXA ETÁRIA <input type="radio"/> 18 A 24 ANOS <input type="radio"/> 25 A 35 ANOS <input type="radio"/> 35 A 45 ANOS <input type="radio"/> ACIMA DE 45 ANOS
6. A natureza da principal atividade que v. desenvolve com o auxílio do PRODUTO é essencialmente de:	<input type="radio"/> NATUREZA DAS ATIVIDADES COM O PRODUTO <input type="radio"/> PESQUISA & DESENVOLVIMENTO <input type="radio"/> APLICAÇÃO IMEDIATA DE UTILITÁRIOS <input type="radio"/> TREINAMENTO & APRENDIZAGEM
7. No contexto de suas atividades v. utiliza o PRODUTO em:	<input type="checkbox"/> UTILIZAÇÃO DO PRODUTO EM <input type="checkbox"/> Construção de Modelos <input type="checkbox"/> Verificação de Modelos Existentes <input type="checkbox"/> Utilização em Teses/Dissertações <input type="checkbox"/> Uso em Trabalhos de Iniciação Científica <input type="checkbox"/> Prestação de Serviços <input type="checkbox"/> Execução de Projetos <input type="checkbox"/> Desenvolvimento de Produtos <input type="checkbox"/> Treinamento em Nível de Extensão <input type="checkbox"/> Suporte a Disciplinas Teóricas (Graduação) <input type="checkbox"/> Suporte a Disciplinas Teóricas (Pós-graduação) <input type="checkbox"/> Suporte Laboratorial (Graduação) <input type="checkbox"/> Suporte Laboratorial (Pós-graduação)
8. Sua plataforma computacional é:	<input type="radio"/> PLATAFORMA COMPUTACIONAL <input type="radio"/> UNIX <input type="radio"/> PC <input type="radio"/> MAC <input type="radio"/> OUTRA
9. Você tem experiência prévia com sistemas computacionais?	<input type="radio"/> EXPERIÊNCIA COMPUTACIONAL PRÉVIA <input type="radio"/> SIM <input type="radio"/> NÃO
10. Há quanto tempo v. usa sistemas computacionais?	<input type="radio"/> TEMPO DE USO DE SISTEMAS COMPUTACIONAIS <input type="radio"/> MENOS DE 3 MESES <input type="radio"/> ENTRE 3 MESES E 1 ANO <input type="radio"/> MAIS DE 1 ANO
11. Com que frequência v. usa sistemas computacionais?	<input type="radio"/> FREQUÊNCIA DE USO DE SISTEMAS COMPUTACIONAIS <input type="radio"/> DIARIAMENTE <input type="radio"/> MAIS DE 1 VEZ POR SEMANA <input type="radio"/> 1 VEZ POR SEMANA <input type="radio"/> MENOS DE 1 VEZ POR SEMANA <input type="radio"/> 1 VEZ POR QUINZENA <input type="radio"/> MENOS DE 1 VEZ POR QUINZENA <input type="radio"/> 1 VEZ POR MÊS <input type="radio"/> ESTOU USANDO PELA 1ª VEZ
12. Você tem experiência prévia com o PRODUTO?	<input type="radio"/> EXPERIÊNCIA PRÉVIA COM O PRODUTO <input type="radio"/> SIM <input type="radio"/> NÃO
13. Há quanto tempo v. usa o PRODUTO?	<input type="radio"/> TEMPO DE USO DO PRODUTO <input type="radio"/> MENOS DE 3 MESES <input type="radio"/> DE 3 MESES A 1 ANO <input type="radio"/> MAIS DE 1 ANO
14. Com que frequência v. usa o PRODUTO?	<input type="radio"/> FREQUÊNCIA DE USO DO PRODUTO <input type="radio"/> DIARIAMENTE <input type="radio"/> MAIS DE 1 VEZ POR SEMANA <input type="radio"/> 1 VEZ POR SEMANA <input type="radio"/> MENOS DE 1 VEZ POR SEMANA <input type="radio"/> 1 VEZ POR QUINZENA <input type="radio"/> MENOS DE 1 VEZ POR QUINZENA <input type="radio"/> 1 VEZ POR MÊS <input type="radio"/> É A 1ª VEZ QUE USO O PRODUTO
15. Qual a versão do PRODUTO que v. utiliza atualmente?	<input type="radio"/> VERSÃO DO PRODUTO UTILIZADA <input type="radio"/> v. 5.3 <input type="radio"/> v. 5.2 <input type="radio"/> v. 5.1 <input type="radio"/> v. 5.0 <input type="radio"/> v. 4x <input type="radio"/> v. 3x <input type="radio"/> v. 2x <input type="radio"/> NUNCA USEI NENHUMA VERSÃO
16. Qual a forma de ajuda do PRODUTO que v. costuma utilizar mais frequentemente?	<input type="radio"/> FORMA DE AJUDA DO PRODUTO MAIS UTILIZADA <input type="radio"/> MANUAL <input type="radio"/> AJUDA DE LINHA DE COMANDO <input type="radio"/> HELPWIN <input type="radio"/> HELP DESK (CD-ROM) <input type="radio"/> HELP DESK (INTERNET) <input type="radio"/> DEMOS <input type="radio"/> CONSULTAS ORAIS A OUTROS USUÁRIOS <input type="radio"/> SUPORTE TÉCNICO (FONE/E-MAIL)
17. Quais as Caixas de Ferramentas (Toolboxes) que v. costuma utilizar mais frequentemente?	<input type="checkbox"/> CAIXAS DE FERRAMENTAS MAIS UTILIZADAS <input type="checkbox"/> Chemometrics <input type="checkbox"/> Communications <input type="checkbox"/> Control System <input type="checkbox"/> Frequency Domain System Identification <input type="checkbox"/> Fuzzy Logic <input type="checkbox"/> Higher-Order Spectral Analysis <input type="checkbox"/> Image Processing <input type="checkbox"/> LMI Control <input type="checkbox"/> Mapping <input type="checkbox"/> Model Predictive Control <input type="checkbox"/> NAG Foundation <input type="checkbox"/> Neural Network <input type="checkbox"/> Optimization <input type="checkbox"/> Partial Differential Equation <input type="checkbox"/> QFT Control Design <input type="checkbox"/> Robust Control <input type="checkbox"/> RTW <input type="checkbox"/> Signal Processing <input type="checkbox"/> Spline <input type="checkbox"/> Statistics <input type="checkbox"/> Symbolic/Extended Symbolic Math <input type="checkbox"/> System Identification <input type="checkbox"/> Wavelet <input type="checkbox"/> μ -Analysis and Synthesis
18. Você tem experiência prévia com produtos similares?	<input type="radio"/> EXPERIÊNCIA PRÉVIA COM PRODUTOS SIMILARES <input type="radio"/> SIM <input type="radio"/> NÃO
19. Se v. respondeu NÃO ao item anterior, encerre neste ponto o preenchimento do questionário. Caso contrário, especifique o produto.	PRODUTO 1
20. Se v. preencheu o item anterior, especifique o tempo de uso do produto.	<input type="radio"/> TEMPO DE USO DO PRODUTO 1 <input type="radio"/> MENOS DE 3 MESES <input type="radio"/> ENTRE 3 MESES E 1 ANO <input type="radio"/> MAIS DE 1 ANO
21. Se v. tem experiência com um segundo produto similar, especifique-o.	PRODUTO 2
22. Se v. preencheu o item anterior, especifique o tempo de uso do produto.	<input type="radio"/> TEMPO DE USO DO PRODUTO 2 <input type="radio"/> MENOS DE 3 MESES <input type="radio"/> ENTRE 3 MESES E 1 ANO <input type="radio"/> MAIS DE 1 ANO
23. Se v. tem experiência com um terceiro produto similar, especifique-o.	PRODUTO 3
24. Se v. preencheu o item anterior, especifique o tempo de uso do produto.	<input type="radio"/> TEMPO DE USO DO PRODUTO 3 <input type="radio"/> MENOS DE 3 MESES <input type="radio"/> ENTRE 3 MESES E 1 ANO <input type="radio"/> MAIS DE 1 ANO

Por favor verifique se preencheu todos os itens do questionário. Só então encerre o processo selecionando com o mouse o botão **OK**. Muito grato por sua colaboração!

OK



Este questionário integrado tem o propósito de coletar informações acerca de como você se sente a respeito do uso do PRODUTO sob condições de teste e de sua documentação *online* e *offline*. Por favor analise cada aspecto considerado, selecionando a opção que mais adequa à sua condição de usuário.

O Produto - Uso e Navegação						
ASPECTO		ESCALA				
		1	2	3	4	5
		Muito fácil	Fácil	Nem fácil nem difícil	Difícil	Muito difícil
1	Uso do produto na realização de tarefas de interesse	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2	Comunicação com o produto (terminologia, linguagem, retorno da informação e das ações em geral)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3	Localização dos itens de menu associados às tarefas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4	Acesso aos itens de menu associados às tarefas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5	Acesso às instruções e advertências do produto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6	Compreensão das instruções e advertências do produto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7	Acesso às janelas de diálogo do produto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8	Compreensão dos termos e solicitações apresentadas nas janelas de diálogo do produto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9	Recuperação de situações de erro	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10	Recuperação de situações de travamento	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11	Compreensão das mensagens de erro apresentadas	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12	Navegação através das diferentes opções do menu, janelas de diálogo e barras de ícones do produto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13	Navegação através de diferentes partes das tarefas executadas pelo produto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14	Memorização das seqüências de ações associadas a cada tarefa executada com o auxílio do produto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15	Aprendizado de novas seqüências de ações	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16	Uso das funcionalidades mais comuns do produto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17	Exploração de novas funcionalidades do produto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18	Visualização à primeira vista das seqüências de ações necessárias para completar as tarefas de interesse com o auxílio do produto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19	Visualização à primeira vista da estruturação dos menus, barras de ícones ou listas de informações disponibilizadas pelo produto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
20	Processo de entrada e saída de dados durante o uso do produto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

O Produto - Documentação <i>Online</i> e <i>Offline</i>						
ASPECTO		ESCALA				
		1	2	3	4	5
		Muito fácil	Fácil	Nem fácil nem difícil	Difícil	Muito difícil
21	Localização e acesso aos mecanismos de <i>ajuda online</i> do produto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
22	Uso dos mecanismos de <i>ajuda online</i> do produto	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
23	Localização e acesso às informações de interesse na <i>ajuda online</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
24	Compreensão das informações de interesse existentes na <i>ajuda online</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25	Linguagem utilizada na <i>ajuda online</i>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
26	Localização das informações de interesse no manual do produto (<i>ajuda offline</i>)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
27	Uso do manual do produto (<i>ajuda offline</i>)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
28	Compreensão das informações de interesse existentes no manual do produto (<i>ajuda offline</i>)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
29	Linguagem utilizada no manual do produto (<i>ajuda offline</i>)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
30	Superação de dúvidas e problemas encontrados durante as sessões de uso do produto a partir dos mecanismos oferecidos tanto pela <i>ajuda online</i> quanto pelo manual do produto (<i>ajuda offline</i>)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Você e o Produto						
ASPECTO		ESCALA				
		1	2	3	4	5
		Concordo totalmente	Concordo parcialmente	Nem concordo nem discordo	Discordo parcialmente	Discordo totalmente
31	Acho que a apresentação do produto é bastante atraente, o que estimula seu uso por mim.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32	Acho que a apresentação do produto facilita o aprendizado rápido de seus comandos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
33	Acho o modo de apresentação das mensagens do produto bastante <u>claro</u> e <u>compreensível</u> .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
34	A resposta do produto às minhas entradas é muito lenta.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
35	Sempre me sinto no controle das ações quando uso o produto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
36	Consigo executar as tarefas de modo <u>direto</u> ao usar o produto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
37	Acho que o produto atende plenamente às minhas necessidades.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
38	Perco muito tempo tentando aprender os comandos do produto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
39	Consigo fazer <u>exatamente</u> o que desejo com os recursos oferecidos pelo produto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
40	As funcionalidades oferecidas pelo produto são facilmente configuráveis de modo a atenderem às minhas necessidades.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
41	O produto apresenta tantas funcionalidades que às vezes acho que nunca aprenderei ou terei necessidade de usá-las todas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
42	Ao usar o produto, consulto frequentemente os mecanismos de ajuda oferecidos.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43	Não acho que as informações da ajuda deste produto são suficientemente <u>eficazes</u> para tirar minhas dúvidas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
44	Ao consultar a ajuda do produto, sempre perco muito tempo tentando encontrar as informações de interesse.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
45	De um modo geral, a <u>quantidade</u> de informações oferecidas pela ajuda do produto é <u>insuficiente</u> para solucionar meus problemas e dúvidas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
46	De um modo geral, a <u>qualidade</u> das informações oferecidas pela ajuda do produto <u>não</u> contribui para a solução dos meus problemas e dúvidas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
47	Sinto-me <u>recioso</u> em alguns momentos das sessões de uso do produto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
48	Em alguns momentos, sinto-me <u>frustrado</u> com o modo como o produto executa tarefas de meu interesse.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
49	De um modo geral, sinto-me <u>satisfeito</u> ao usar o produto.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
50	Recomendaria <u>sem hesitação</u> o uso do produto aos meus colegas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

ANEXO I

Ficha Cadastral de Participação em Ensaios de Usabilidade



Universidade Federal da Paraíba
Centro de Ciências e Tecnologia
Departamento de Engenharia Elétrica
Grupo de Interfaces Homem-Máquina



DECLARAÇÃO DAS CONDIÇÕES **DE REALIZAÇÃO DO ENSAIO**

Prezado Participante,

Você está sendo convidado a participar da avaliação do processo interativo do **MATLAB**, um produto desenvolvido pela **The Math Works Inc.**, cuja etapa de coleta de informações foi planejada para ocorrer em 03 (três) estágios distintos, a saber:

- Inicialmente, v. preencherá um instrumento de sondagem do seu perfil de usuário do produto, um questionário *online* denominado **DePerUSI** (*Delineamento do Perfil do Usuário de Sistemas Interativos*);
- Em seguida, v. participará de uma sessão de testes do produto, mediada por **Eustáquio Rangel** e **Fátima Turnell**, que estarão à sua disposição para esclarecer quaisquer dúvidas e detalhes relativos ao procedimento avaliatório;
- Finalmente, v. preencherá um segundo instrumento de sondagem, denominado **OpUS** (*Opinião do Usuário de Software*), destinado à coleta de informações sobre como v. se sente a respeito do uso do produto e de sua documentação *online* (*helpwin*) e *offline* (*manuais*).

Durante a segundo estágio da etapa de coleta de informações, v. será instruído a realizar uma série de tarefas com o auxílio do **MATLAB**, que nos possibilitarão a aquisição de informações valiosas para a realização posterior de análises visando a otimização do processo interativo usuário-produto.

É importante ter em mente que o alvo do procedimento avaliatório **não** será v. e sim o produto em questão. Cada sessão de testes será de **60 minutos**, durante os quais serão feitos registros escritos e áudio-visuais de detalhes pertinentes ao contexto da avaliação. É importante que fique bastante claro que tais registros serão empregados **apenas** para fins de avaliação e otimização do sistema, **não** devendo ser distribuídos, consultados por indivíduos alheios ao processo de avaliação ou apresentados de alguma forma que possibilite a sua identificação enquanto participante deste ensaio.

Seus direitos na qualidade de participante deste ensaio são:

1. Poder solicitar a exclusão de sua participação nos testes em qualquer instante e por quaisquer razões que v. julgar convenientes;
2. Poder verificar ao final da sessão de testes, caso julgue necessário, os dados coletados a partir de sua interação com o produto. Caso decida excluí-los do contexto do ensaio avaliatório, v. deverá informar imediatamente sua decisão aos avaliadores. Caso contrário, tenha a certeza de que nos comprometemos com a inviabilização de qualquer tentativa de identificação desses dados por terceiros;
3. Esclarecer dúvidas e questionar detalhes referentes ao ensaio, porém apenas quando indicado pelos avaliadores. Caso os esclarecimentos solicitados venham a comprometer a integridade dos dados ou polarizar, de algum modo, sua opinião, v. será informado pelos avaliadores, que poderão se omitir em respondê-los.

Solicitamos que seja evitada qualquer discussão desta sessão com outros indivíduos, pertencentes ou não ao grupo de usuários de teste deste ensaio.

Por fim, desejamos agradecer-lhe pelo dispêndio de tempo e esforço, assim como pela a gentileza em aceitar nosso convite para participar deste ensaio. Esteja certo de que neste processo não há respostas **certas** ou **erradas**, nem estão envolvidos conceitos de **bom** e **mau desempenho**. Insistimos em reforçar o fato de que a sessão de testes da qual v. participará objetiva tão somente a identificação de problemas de usabilidade associados à sua interação com o produto!

Se ainda houver questões relativas ao teor do ensaio ou de seu engajamento no processo avaliatório, contate-nos nos ramais **1357** ou **1143** ou envie-nos um *e-mail* para os endereços **rangel@dee.ufpb.br** ou **turnellm@dee.ufpb.br**. Teremos prazer em respondê-las.

Enviamos em anexo uma ficha de cadastro, cujo preenchimento confirmará sua concordância total e voluntária com os termos deste documento e, adicionalmente, o tornará um participante em potencial de pesquisas futuras do **Grupo de Interfaces Homem-Máquina** do DEE/CCT/UFPB.

Atenciosamente,

Maria de Fátima Queiroz Vieira Turnell

Chefe do GIHM/DEE/CCT/UFPB

José Eustáquio Rangel de Queiroz

GIHM/DEE/CCT/UFPB – Doutorando



ANEXO J

J1 – Roteiro das Tarefas de Teste

Tarefa 0 (Pré-Tarefa)

TAREFA 0 (PRÉ-TAREFA): Inicialização do sistema, criação de pasta de trabalho e ativação do aplicativo

Tempo Estimado: 05 min.

Roteiro: Você irá participar da execução de algumas atividades que envolvem o uso do **MATLAB**. Uma das metas dessas atividades é a entrada, o processamento e a saída de dados numéricos e gráficos utilizando o **MATLAB** como ferramenta de suporte. Antes de qualquer ação *realmente* destinada ao desenvolvimento de tais atividades, você deverá começar sua sessão de trabalho inicializando o sistema, criando um espaço para armazenamento de seus dados e ativando o **MATLAB**, como você o faz usualmente, a fim de poder realizar as próximas etapas do trabalho.

Instruções:

- Ligue o sistema computacional e aguarde a inicialização do *Windows*.
- Crie uma pasta de trabalho pessoal em **E:\USERS\ENSAIO**.
- Transfira todos os arquivos da pasta **E:\USERS\RANGEL\ENSAIO** para sua pasta de trabalho, onde também deverão ser salvos todos os arquivos que você criar durante a sessão.
- Abra o **MATLAB v. 5.3**.
- Passe para a etapa seguinte desta sessão de trabalho, descrita na próxima página.

ANEXO J

J2 – Roteiro das Tarefas de Teste

Tarefa 1

TAREFA 1: Criação de arquivos de dados

Tempo Estimado: **10 min.**

Roteiro: No contexto das atividades que você está prestes a iniciar envolvendo o uso do **MATLAB**, você trabalhará com conjuntos de dados numéricos ainda não armazenados em arquivos. Portanto, você deverá criar arquivos de dados, conforme as instruções apresentadas a seguir.

Instruções:

- ❑ Crie um novo arquivo para armazenamento dos dados que serão processados nas etapas seguintes desta sessão de trabalho.
- ❑ Digite em uma única coluna o conjunto numérico apresentado a seguir, mantendo a ordem crescente de arranjo dos dados.

0	6,94
0,23	7,02
0,51	7,18
0,92	7,57
1,01	7,78
1,32	7,95
1,46	8,14
1,75	8,35
1,89	8,78
2,03	8,89
2,11	9,01
2,45	9,17
2,88	9,21
3,31	9,38
3,70	9,50
4,12	9,75
4,65	9,83
4,90	9,89
5,12	9,92
5,30	9,95
5,55	9,98
5,80	10,00
5,98	10,20
6,15	10,89
6,36	10,99
6,55	11,18
6,71	11,32
6,85	12,00

- ❑ **Salve o arquivo com o nome *dados1.dat*.**
- ❑ **Crie um outro arquivo para armazenar o seguinte conjunto numérico:**

0,0000
0,2500
0,5000
0,7500
1,0000
1,2500
1,5000
1,7500
2,0000
2,2500
2,5000
2,7500
3,0000
3,2500
3,5000

- ❑ **Salve o arquivo com o nome *dados2.dat***
- ❑ **Passa para a TAREFA 2, descrita na próxima página.**

ANEXO J

J3 – Roteiro das Tarefas de Teste

Tarefa 2

TAREFA 2: Execução, abertura, e correção de *script*

Tempo Estimado: **10 min.**

Roteiro: Após o armazenamento do conjunto de dados numéricos realizado na tarefa anterior, você deverá executar um *script* do **MATLAB v. 5.3** e, dependendo dos resultados obtidos, abri-lo e corrigi-lo, procedendo conforme as instruções abaixo.

Instruções:

- Carregue o arquivo ***dados1.dat***.
- Execute o *script* ***teste1.m*** e verifique os resultados apresentados pelo aplicativo.
- Se o **MATLAB v. 5.3** não apontar erros, passe para a etapa seguinte desta sessão de trabalho, descrita na próxima página.
- Em caso contrário, o **MATLAB v. 5.3** emitirá tantas mensagens quantos forem os erros assinalados pelo aplicativo durante a execução do *script* ***teste1***.
- Tente interpretar cada mensagem de erro, a fim de identificar a natureza de cada erro e corrigi-lo.
- De acordo com a natureza de cada erro apontado, abra o *script* ***teste1.m*** e/ou o arquivo ***dados1.dat*** e corrija o que estiver incorreto.
- Execute novamente o *script* ***teste1.dat***
- Se o **MATLAB v. 5.3** não apontar nenhum outro erro, passe para a próxima página.
- Em caso contrário, abra novamente o *script* ***teste1.m*** e/ou o arquivo ***dados1.dat***, corrigindo o que **AINDA** estiver incorreto.
- Execute ***teste1*** pela última vez. Desta vez, mesmo que **AINDA** haja erros, apenas comente verbalmente, se achar necessário, os resultados obtidos.
- Passe para a próxima etapa desta sessão de trabalho, descrita a seguir.

ANEXO J

J4 – Roteiro das Tarefas de Teste

Tarefa 3

TAREFA 3: Alteração de dados numéricos

Tempo Estimado: **10 min.**

Roteiro: Nesta tarefa, um conjunto de dados numéricos anteriormente armazenado deverá ser alterado. Siga as instruções abaixo, a fim de alterá-lo convenientemente.

Instruções:

- Carregue o arquivo *dados2.dat*.
- Execute o *script teste2.m*. Os dados resultantes não deverão apresentar funções descontínuas.
- Analise os resultados numéricos e gráficos obtidos, verificando se há erros de digitação e/ou valores responsáveis por descontinuidades na função definida em *teste2.m*.
- Elimine os erros identificados.
- Execute novamente o *script teste2.m* e verifique se os resultados foram alterados a contento.
- Mais uma vez, apenas comente verbalmente os resultados obtidos, passando para a próxima etapa desta sessão de trabalho, descrita na página seguinte.

ANEXO J

J5 – Roteiro das Tarefas de Teste

Tarefa 4

TAREFA 4: Configuração de representações gráficas e visualização em tela

Tempo Estimado: **10 min.**

Roteiro: No contexto das atividades que você tem desenvolvido com o **MATLAB**, desde o início desta sessão, você deverá alterar, nesta etapa, tanto alguns dos atributos de visualização da representação gráfica associada à execução do *script teste3.m*, quanto alguns dos atributos relativos à impressão desta representação. Faça-o segundo as instruções apresentadas a seguir.

Instruções:

- ❑ Execute o *script teste3.m* existente na pasta **E:\USERS\RANGEL\ENSAIO\AUXI**.
- ❑ Configure a visualização dos gráficos visualizados na tela de modo a atender às especificações apresentadas a seguir.

CURVA	COMPOSIÇÃO CROMÁTICA			Espessura	Marcador	Estilo
	Vermelho	Verde	Azul			
1	218	13	213	2	o	NENHUM
2	72	102	45	2	NENHUM	sólido
3	219	36	137	0,5	NENHUM	pontilhada

- **Tamanho do marcador: 5**
- **Texto: Arial 11 NEGRITO**

PLANO DO GRÁFICO:

- **Azimute: -45°**
- **Elevação: 60°**
- ❑ **Em seguida, configure a página para atender aos seguintes requisitos:**
 - **Orientação: RETRATO**
 - **Impressão: Colorida**
 - **Posição: MÁXIMA RELAÇÃO DE ASPECTO**

- Em seguida, configure o processo de impressão para atender aos seguintes requisitos:
 - Tamanho do papel: **A4**
 - Orientação do papel: **PAISAGEM**
 - Qualidade da impressão: **MELHOR**
- Passe, então, para a próxima e última etapa desta sessão de trabalho, descrita na próxima página.

The following information is for your information only. It is not intended to be used as a substitute for professional advice. Please consult your attorney for more information.

STATE OF CALIFORNIA
DEPARTMENT OF REVENUE
SALES TAX

Period: 1/1/2018 to 12/31/2018
Total Sales Tax: \$1,234,567.89
Total Refunds: \$123,456.78
Net Sales Tax: \$1,111,111.11



ANEXO K

K1 – Roteiro das Tarefas de Teste

Tarefa 0 (Pré-Tarefa)

TAREFA 0 (PRÉ-TAREFA): Inicialização do sistema, criação de pasta de trabalho e ativação do aplicativo

Objetivo Geral: Reprodução de um início de sessão em ambientes de trabalho típicos, onde a primeira ação do usuário é a inicialização do sistema e do aplicativo com o qual trabalha;

Objetivos Específicos:

1. Observação da *facilidade de execução da tarefa*;
2. Mensuração do *tempo de conclusão da tarefa*;
3. Mensuração do *número de ações incorretas* durante a execução da tarefa;
4. Mensuração do *número de escolhas incorretas* durante a execução da tarefa;
5. Mensuração do *número de erros repetidos* durante a execução da tarefa;

Indicadores:

1. facilidade de execução da tarefa (*observação direta*)
2. tempo de conclusão da tarefa (*mensuração do desempenho*)
3. número de ações incorretas (*mensuração do desempenho*)
4. número de escolhas incorretas (*mensuração do desempenho*)
5. número de erros repetidos (*mensuração do desempenho*)

Tempo Estimado: 05 min.

Roteiro: Você irá participar da execução de algumas atividades que envolvem o uso do **MATLAB**. Uma das metas dessas atividades é a entrada, o processamento e a saída de dados numéricos e gráficos utilizando o **MATLAB** como ferramenta de suporte. Antes de qualquer ação *realmente* destinada ao desenvolvimento de tais atividades, você deverá começar sua sessão de trabalho inicializando o sistema, criando um espaço para armazenamento de seus dados e ativando o **MATLAB**, como você o faz usualmente, a fim de poder realizar as próximas etapas do trabalho.

Instruções:

- ❑ Ligue o sistema computacional e aguarde a inicialização do *Windows*.
- ❑ Crie uma pasta de trabalho pessoal em **E:\USERS\ENSAIO**.
- ❑ Transfira todos os arquivos da pasta **E:\USERS\RANGEL\ENSAIO** para sua pasta de trabalho, onde também deverão ser salvos todos os arquivos que você criar durante a sessão.
- ❑ Abra o **MATLAB v. 5.3**.
- ❑ Passe para a etapa seguinte desta sessão de trabalho, descrita na próxima página.

ANEXO K

K2 – Roteiro das Tarefas de Teste

Tarefa 1

TAREFA 1: Criação de arquivos de dados

Objetivo Geral: Detecção de problemas no processo de entrada de dados;

- Objetivos Específicos:**
1. Observação da *facilidade de uso do produto*;
 2. Observação da *facilidade de execução da tarefa utilizando o MATLAB v. 5.3*;
 3. Mensuração do *tempo de conclusão da tarefa*;
 4. Mensuração do *número de ações incorretas* durante a execução da tarefa;
 5. Mensuração do *número de escolhas incorretas* durante a execução da tarefa;
 6. Mensuração do *número de erros repetidos* durante a execução da tarefa;

- Indicadores:**
1. facilidade de execução da tarefa (*observação direta*)
 2. tempo de conclusão da tarefa (*mensuração do desempenho*)
 3. número de ações incorretas (*mensuração do desempenho*)
 4. número de escolhas incorretas (*mensuração do desempenho*)
 5. número de erros repetidos (*mensuração do desempenho*)

Tempo Estimado: 10 min.

Roteiro: No contexto das atividades que você está prestes a iniciar envolvendo o uso do **MATLAB**, você trabalhará com conjuntos de dados numéricos ainda não armazenados em arquivos. Portanto, você deverá criar arquivos de dados, conforme as instruções apresentadas a seguir.

Instruções:

- Crie um novo arquivo para armazenamento dos dados que serão processados nas etapas seguintes desta sessão de trabalho.

- ❑ Digite em uma única coluna o conjunto numérico apresentado a seguir, mantendo a ordem crescente de arranjo dos dados.

0	6,94
0,23	7,02
0,51	7,18
0,92	7,57
1,01	7,78
1,32	7,95
1,46	8,14
1,75	8,35
1,89	8,78
2,03	8,89
2,11	9,01
2,45	9,17
2,88	9,21
3,31	9,38
3,70	9,50
4,12	9,75
4,65	9,83
4,90	9,89
5,12	9,92
5,30	9,95
5,55	9,98
5,80	10,00
5,98	10,20
6,15	10,89
6,36	10,99
6,55	11,18
6,71	11,32
6,85	12,00

- ❑ Salve o arquivo com o nome *dados1.dat*.
- ❑ Crie um outro arquivo para armazenar o seguinte conjunto numérico:

0,0000
0,2500
0,5000
0,7500
1,0000
1,2500
1,5000
1,7500
2,0000
2,2500
2,5000
2,7500
3,0000
3,2500
3,5000

- ❑ Salve o arquivo com o nome *dados2.dat*
- ❑ Passe para a TAREFA 2, descrita na próxima página.

ANEXO K

K3 – Roteiro das Tarefas de Teste

Tarefa 2

TAREFA 2: Execução, abertura, e correção de script

Objetivo Geral: Detecção de problemas no processo de manipulação de dados;

- Objetivos Específicos:**
1. Observação da *facilidade de uso do produto*;
 2. Observação da *facilidade de execução da tarefa utilizando o MATLAB v. 5.3*;
 3. Mensuração do *tempo de conclusão da tarefa*;
 4. Mensuração do *número de ações incorretas durante a execução da tarefa*;
 5. Mensuração do *número de escolhas incorretas durante a execução da tarefa*;
 6. Mensuração do *número de erros repetidos durante a execução da tarefa*;

- Indicadores:**
1. facilidade de execução da tarefa (*observação direta*)
 2. tempo de conclusão da tarefa (*mensuração do desempenho*)
 3. número de ações incorretas (*mensuração do desempenho*)
 4. número de escolhas incorretas (*mensuração do desempenho*)
 5. número de erros repetidos (*mensuração do desempenho*)

Tempo Estimado: 10 min.

Roteiro: Após o armazenamento do conjunto de dados numéricos realizado na tarefa anterior, você deverá executar um *script* do *MATLAB v. 5.3* e, dependendo dos resultados obtidos, abri-lo e corrigi-lo, procedendo conforme as instruções abaixo.

Instruções:

- Carregue o arquivo *dados1.dat*.
- Execute o *script teste 1.m* e verifique os resultados apresentados pelo aplicativo.
- Se o *MATLAB v. 5.3* não apontar erros, passe para a etapa seguinte desta sessão de trabalho, descrita na próxima página.
- Em caso contrário, o *MATLAB v. 5.3* emitirá tantas mensagens quantos forem os

erros assinalados pelo aplicativo durante a execução do *script teste 1*.

- ❑ Tente interpretar cada mensagem de erro, a fim de identificar a natureza de cada erro e corrigi-lo.
- ❑ De acordo com a natureza de cada erro apontado, abra o *script teste 1.m* e/ou o arquivo *dados 1.dat* e corrija o que estiver incorreto.
- ❑ Execute novamente o *script teste 1.dat*
- ❑ Se o *MATLAB v. 5.3* não apontar nenhum outro erro, passe para a próxima página.
- ❑ Em caso contrário, abra novamente o *script teste 1.m* e/ou o arquivo *dados 1.dat*, corrigindo o que AINDA estiver incorreto.
- ❑ Execute *teste 1* pela última vez. Desta vez, mesmo que AINDA haja erros, apenas comente verbalmente, se achar necessário, os resultados obtidos.
- ❑ Passe para a próxima etapa desta sessão de trabalho, descrita a seguir.

ANEXO K

K4 – Roteiro das Tarefas de Teste

Tarefa 3

TAREFA 3: Alteração de dados numéricos

Objetivo Geral: Detecção de problemas no processo de manipulação de dados;

- Objetivos Específicos:**
1. Observação da *facilidade de uso do produto*;
 2. Observação da *facilidade de execução da tarefa utilizando o MATLAB v. 5.3*;
 3. Mensuração do *tempo de conclusão da tarefa*;
 4. Mensuração do *número de ações incorretas* durante a execução da tarefa;
 5. Mensuração do *número de escolhas incorretas* durante a execução da tarefa;
 6. Mensuração do *número de erros repetidos* durante a execução da tarefa;

- Indicadores:**
1. facilidade de execução da tarefa (*observação direta*)
 2. tempo de conclusão da tarefa (*mensuração do desempenho*)
 3. número de ações incorretas (*mensuração do desempenho*)
 4. número de escolhas incorretas (*mensuração do desempenho*)
 5. número de erros repetidos (*mensuração do desempenho*)

Tempo Estimado: 10 min.

Roteiro: Nesta tarefa, um conjunto de dados numéricos anteriormente armazenado deverá ser alterado. Siga as instruções abaixo, a fim de alterá-lo convenientemente.

Instruções:

- ❑ Carregue o arquivo *dados2.dat*.
- ❑ Execute o *script teste2.m*. Os dados resultantes não deverão apresentar funções descontínuas.
- ❑ Analise os resultados numéricos e gráficos obtidos, verificando se há erros de digitação e/ou valores responsáveis por descontinuidades na função definida em *teste2.m*.

- ❑ **Elimine os erros identificados.**
- ❑ **Execute novamente o *script teste2.m* e verifique se os resultados foram alterados a contento.**
- ❑ **Mais uma vez, apenas comente verbalmente os resultados obtidos, passando para a próxima etapa desta sessão de trabalho, descrita na página seguinte.**

ANEXO K

K5 – Roteiro das Tarefas de Teste

Tarefa 4

TAREFA 4: Configuração de representações gráficas e visualização em tela

Objetivo Geral: Detecção de problemas no processo de configuração da saída de dados;

- Objetivos Específicos:**
1. Observação da *facilidade de uso do produto*;
 2. Observação da *facilidade de execução da tarefa utilizando o MATLAB v. 5.3*;
 3. Mensuração do *tempo de conclusão da tarefa*;
 4. Mensuração do *número de ações incorretas durante a execução da tarefa*;
 5. Mensuração do *número de escolhas incorretas durante a execução da tarefa*;
 6. Mensuração do *número de erros repetidos durante a execução da tarefa*;

- Indicadores:**
1. facilidade de execução da tarefa (*observação direta*)
 2. tempo de conclusão da tarefa (*mensuração do desempenho*)
 3. número de ações incorretas (*mensuração do desempenho*)
 4. número de escolhas incorretas (*mensuração do desempenho*)
 5. número de erros repetidos (*mensuração do desempenho*)

Tempo Estimado: 10 min.

Roteiro: No contexto das atividades que você tem desenvolvido com o **MATLAB**, desde o início desta sessão, você deverá alterar, nesta etapa, tanto alguns dos atributos de visualização da representação gráfica associada à execução do *script teste3.m*, quanto alguns dos atributos relativos à impressão desta representação. Faça-o segundo as instruções apresentadas a seguir.

Instruções:

- ❑ Execute o *script teste3.m* existente na pasta **E:\USERS\RANGEL\ENSAIO\AUXI**.
- ❑ Configure a visualização dos gráficos visualizados na tela de modo a atender às especificações apresentadas a seguir.

CURVA	COMPOSIÇÃO CROMÁTICA			Espessura	Marcador	Estilo
	Vermelho	Verde	Azul			
1	218	13	213	2	0	NENHUM
2	72	102	45	2	NENHUM	sólido
3	219	36	137	0,5	NENHUM	pontilhada

- **Tamanho do marcador: 5**
- **Texto: Arial 11 NEGRITO**

PLANO DO GRÁFICO:





- **Azimute: -45°**
 - **Elevação: 60°**
- **Em seguida, configure a página para atender aos seguintes requisitos:**
- **Orientação: RETRATO**
 - **Impressão: Colorida**
 - **Posição: MÁXIMA RELAÇÃO DE ASPECTO**
- **Em seguida, configure o processo de impressão para atender aos seguintes requisitos:**
- **Tamanho do papel: A4**
 - **Orientação do papel: Paisagem**
 - **Qualidade da impressão: Melhor**
- **Passa, então, para a próxima e última etapa desta sessão de trabalho, descrita na próxima página.**

ANEXO L

L1 – Ficha de Registro de Eventos - Tarefas Individuais

Produto:			Participante:		
Registro de Eventos de Teste - TAREFA _____					
					
LEGENDA			 Tempo de Leitura	 Tempo de Execução	
 Nº de Ações Incorretas	 Nº de Escolhas Incorretas	 Nº de Erros Repetidos			
EVENTO			COMENTÁRIO		
1					
					
2					
					
3					
					
4					
					
6					
					
6					
					
7					
					
8					
					
9					
					
10					
					

L2 – Ficha de Registro de Eventos - Sumário de Sessão

Produto:		Participante:	Data da Sessão: ____/____/____			
			Início da Sessão: ____:____:____			
			Término da Sessão: ____:____:____			
Tarefa	Descrição da Tarefa	Notas de Interesse				
Notas Adicionais:						

ANEXO M

Tópicos de Estatística Indutiva

O propósito desta revisão básica de Estatística Indutiva é respaldar a discussão dos resultados da mensuração do desempenho e da sondagem da satisfação do usuário, apresentados nas seções 7.3 e 7.4, respectivamente. Portanto, é necessário ter em mente que a revisão que será apresentada a seguir tratará *apenas* de conceitos e procedimentos que foram empregados no processamento dos dados quantitativos coletados a partir do teste de usabilidade, assim como dos dados qualitativos obtidos a partir da aplicação do instrumento de sondagem *OpUS*, descrito no capítulo anterior.

M1 Noções Conceituais

Na pesquisa científica, raramente se coleta informações de todos os membros de uma *população* para que se possa fazer declarações confiáveis sobre características ou atributos daquela população, as quais assumem diferentes valores para os elementos que a compõem (*variáveis*). Quer por conta da impossibilidade ou de dificuldades das mais diversas naturezas, quer por não ser necessário um *censo populacional*¹, costuma-se efetuar medidas apenas sobre um subconjunto da população total, denominado *amostra* [Burt96]. Graças aos métodos concebidos no âmbito da Estatística Indutiva, também conhecida como Inferência Estatística, os resultados obtidos a partir de um processo de *amostragem* podem ser convenientemente processados, fundamentando estimativas de determinadas características de uma população (estudos *analíticos*) ou tomadas de decisões referentes àquela população (estudos *enumerativos*) com base apenas em resultados de amostras [Levi98].

Evidentemente, os processos de amostragem apresentam a desvantagem de restringir a atenção do pesquisador a uma pequena proporção da população investigada, impossibilitando-lhe de ser tão acurado sobre as características consideradas da população quanto o seria se realizasse um censo completo [Burt96]. O processo se torna mais suscetível à ocorrência de erros (e.g., *de abrangência, por falta de resposta, de amostragem ou de medição*).

Entretanto, é possível atingir bons níveis de acurácia e confiabilidade nos resultados de uma pesquisa científica através da definição da(s) amostra(s) da população investigada. Quanto mais representativa da população for a amostra selecionada para estudo, mais elevados serão os níveis de acurácia e confiabilidade dos resultados. Dito de outro modo, no que concerne à(s) variável(eis) de interesse da pesquisa, quanto mais próximas das características básicas da população forem as características da amostra, respeitadas as discrepâncias inerentes à

¹ Tabulação completa de uma variável ou de um grupo de variáveis de interesse para todos os elementos de uma população.

aleatoriedade sempre presente, mais acurados serão os resultados obtidos [Coch77, Burt96, Levi98].

Estatisticamente falando, uma *população* (ou *universo*) é a totalidade das observações consideradas em um contexto de estudo, da qual pode ser extraído para análise um número limitado de observações da mesma natureza, denominado *amostra* (ou *universo amostral*), através de um processo de seleção, que pode ser probabilístico ou não probabilístico, denominado *amostragem* [Corm74, Burt96, Levi98, Troc99]. O processo de amostragem será dito *probabilístico* se todos os elementos da população forem selecionados com base em probabilidades conhecidas e não nulas [Coch77, Levi98], i.e. se a probabilidade de qualquer elemento da população pertencente à amostra puder ser determinada [Burt96]. Caso contrário, o processo de amostragem será dito *não probabilístico*.

Relevando sutis diferenças encontradas na literatura sobre classificação de amostras e integrando as classificações consideradas por Trochim [Troc99] e Burton e Barber [Burt96], por se julgar a integração mais adequada a esta fundamentação revisiva, é possível encontrar no campo da pesquisa científica cinco tipos de amostras probabilísticas e quatro tipos de amostras não probabilísticas, conforme discrimina sumariamente o Quadro 41.

Quadro 41 – Caracterização de Amostras Probabilísticas e Não Probabilísticas.

AMOSTRAGEM	TIPOS DE AMOSTRA	CARACTERIZAÇÃO
Probabilística	Aleatória Simples ^{1, 2} (<i>Simple random sample</i>)	Cada elemento ou item da amostra possui a mesma probabilidade de seleção que os demais elementos ou itens da população considerada (<i>populações finitas</i>) ou todas as observações selecionadas para a composição da amostra são estatisticamente independentes (<i>populações finitas</i>).
	Aleatória Sistemática ^{1, 2} (<i>Systematic random sample</i>)	Cada k-ésimo elemento ou item de uma tabela de amostragem da população considerada é selecionado para a composição da amostra.
	Aleatória Estratificada ^{1, 2} (<i>Stratified random sample</i>)	Cada elemento ou item da amostra possui a mesma probabilidade de seleção que os demais elementos ou itens de estratos homogêneos da população considerada e cada estrato também possui a mesma probabilidade de seleção que os demais estratos da população.
	Por agrupamento ou De grupo ^{1, 2} (<i>Cluster sample</i>)	Todos os elementos ou itens pertencentes a estratos definidos por conveniência em uma população estratificada são selecionados para a composição da amostra.
	Adaptativa ou Híbrida ¹ (<i>Multi-stage sample</i>)	Os elementos ou itens que compõem a amostra são selecionados a partir da combinação de duas ou mais das estratégias adotadas para a composição dos tipos de amostras probabilísticas acima discriminados.
Não probabilística	Por julgamento ^{1, 2} (<i>Judgemental or Purposive sample</i>)	Todos os elementos ou itens da população aprovados por julgamento pessoal entram na composição da amostra.
	Casual ou Por conveniência ^{1, 2} (<i>Haphazard, Accidental, Convenience or Accessibility sample</i>)	Apenas os elementos ou itens da população convenientes ou acessíveis são selecionados para a composição da amostra.
	Por quotização ^{1, 2} (<i>Quote sample</i>)	Elementos ou itens de quotas específicas de uma população dividida em quotas fixas com base em uma dada característica de interesse (e.g. sexo, faixa etária, etnia) são selecionados para a composição da amostra.
	Voluntária ^{1, 2} (<i>Volunteer sample</i>)	Indivíduos pertencentes à população que se oferecem como voluntários são agrupados para composição da amostra.

¹ Conforme Trochim [Troc99]

² Conforme Burt e Barber [Burt96]

* Trochim [Troc99] considera a **amostragem por quotização** uma sub-categoria da **amostragem por julgamento**, subdividindo-a ainda em **proporcional** e **não proporcional**, quando mantém ou não, respectivamente, na amostra e na população a mesma proporção da característica na qual se baseou o processo de quotização (e.g., se a característica de interesse for o **sexo** dos indivíduos e se for observada na população a proporção de 4 para 6 entre indivíduos dos sexos feminino e masculino, respectivamente, a amostra por quotização será *proporcional* se seus indivíduos mantiverem esta proporção e *não proporcional* caso contrário).

Embora as técnicas da Estatística Indutiva pressuponham que as amostras utilizadas sejam probabilísticas, em diversos contextos práticos não é possível satisfazer tal pressuposto. No entanto, o bom senso irá indicar quando o processo de amostragem, embora não sendo probabilístico, poderá ser, para efeitos práticos, considerado como tal, de modo que se possa usufruir das possibilidades de utilização do método estatístico em geral [Fech00].

Outra questão relevante no âmbito da Estatística Indutiva é a delimitação do *tamanho* da amostra, tanto no que diz respeito à generalização das conclusões quanto no que concerne às restrições de cunho prático (e.g., disponibilidade de indivíduos, custos, tempo). No tocante ao último aspecto, considere-se como base para a delimitação da amostra a *lei empírica do acaso* ou *lei dos grandes números*, que consagra o princípio de que a aproximação relativa aumenta à medida que cresce o número de determinações. Considere-se também que diversos contextos da pesquisa científica estão relacionados com medições ou sondagens envolvendo populações de tamanhos consideráveis. Por fim, considere-se, por exemplo, que os indivíduos pertencentes a essas populações se encontram, no mundo real, distribuídos em diferentes regiões geográficas ou, caso se encontrem em uma mesma região geográfica, só estejam disponíveis mediante remuneração ou não disponham de tempo para a participação em um experimento estatístico.

Estes dois exemplos, envolvendo explicitamente as variáveis *distribuição geográfica*, *custos* e *tempo* e implicitamente diversas outras variáveis (e.g. *particularidades sócio-culturais*), são suficientes para mostrar que o uso da *lei empírica do acaso* como base para a determinação do tamanho do universo amostral é inadequado em inúmeras situações cotidianas, face às restrições impostas pelo mundo real. Esta constatação conduz a outra, de carácter mais abrangente: na prática, a determinação do tamanho apropriado de uma amostra não é um procedimento trivial, por estar sujeito a restrições de diferentes naturezas.

Assim, parte-se do princípio que uma amostra deve incluir um número suficiente de casos, escolhidos aleatoriamente, que lhe possibilite oferecer uma segurança estatística aceitável em relação à representatividade dos dados. Para tanto, pode-se adotar a recomendação de Pearson, que delimita em 20 o número mínimo de observações ou do tamanho da amostra [Coch77]. Outros autores (e.g. Sabo [Sabo99]) mencionam o limite mínimo de 30 observações. Como regra geral, considera-se como *pequenas amostras* os universos amostrais cujas dimensões não ultrapassam 30 unidades amostrais ($n_A < 30$) e como *grandes amostras* aqueles cujas dimensões excedem o referido montante ($n_A > 30$) [Sabo99, Fech00].

Na pesquisa científica fundamentada em resultados obtidos a partir de processos de amostragem, é imprescindível a comprovação de que os resultados da amostra reflitam os resultados da população, i.e. sejam significativos e apresentem um erro desprezível ou, no pior dos casos, aceitável para o contexto considerado. Isto é possível graças à Teoria das Probabilidades, a partir da qual se averigua a probabilidade de que os resultados obtidos através de amostras reflitam os resultados das populações das quais as referidas amostras foram extraídas.

No âmbito da Estatística Indutiva, um *parâmetro* é uma medida empregada para descrever uma característica comum aos elementos de uma população, enquanto uma *estatística* é uma medida empregada para descrever uma característica comum apenas aos elementos de uma

amostra da população (e.g., a *média de uma população*² é um parâmetro, enquanto a *média de uma amostra*² é uma estatística). Ao calcular parâmetros (e.g., a média de uma população) com base em amostras, as medidas encontradas estão sujeitas a erros. Como já foi anteriormente ressaltado, é de fundamental importância que os erros cometidos em processos dessa natureza sejam desprezíveis ou, na pior das hipóteses, aceitáveis. Uma medida utilizada para avaliar erros de inferência denomina-se *erro padrão*.

O *erro padrão* de uma estatística é o desvio padrão da distribuição amostral daquela estatística. Erros padrões são importantes por refletirem a variabilidade amostral de uma estatística. O erro padrão de uma estatística é função do tamanho da amostra considerada, assim como de sua variabilidade amostral. Em geral, quanto maior o tamanho da amostra menor o erro padrão. O erro padrão de uma estatística é usualmente representado pela letra grega *sigma* (σ), acompanhada de uma subscrição que indica a estatística (e.g., o erro padrão da média é indicado pelo símbolo σ_μ). A Estatística Indutiva envolvida na construção de intervalos de confiança e testes de significância se fundamenta em erros padrões.

M2 Erro Padrão da Média

Conforme mencionado anteriormente, o *erro padrão* (σ) depende do número de observações da amostra considerada, assim como da variabilidade das medidas. Para amostras probabilísticas ou praticamente probabilísticas, o *erro padrão da média* (σ_μ) pode ser facilmente obtido a partir da expressão:

$$\sigma_\mu = \frac{s}{\sqrt{n_A}} \quad (1)$$

onde s representa o *desvio padrão*³ da amostra e n_A o *tamanho* da amostra [Burt96, Levi98, Troc99].

Assim, quanto maior o tamanho da amostra e menor o desvio padrão, tanto menor será o erro padrão da média. Se o erro padrão obtido da amostra for desprezível, ou seja, muito menor do que a média encontrada, a média observada poderá ser tomada como representativa da população estudada, i.e. poderá ser considerada significativa.

Conforme exposto acima, é provável que a média de uma amostra não seja precisamente igual àquela relativa à população. A extensão da discrepância depende do tamanho e da variabilidade da amostra considerada. Se a amostra for grande e pouco dispersa, a amostra média provavelmente será bastante próxima da média populacional. Entretanto, se a amostra for pequena e apresentar uma variabilidade elevada, a média da amostra poderá ser bastante diferente da média populacional. A determinação do *erro padrão* permite a estimativa da média aritmética da população (μ) a partir de dois métodos, a saber: (i) *por ponto*, segundo o qual a estimativa é feita com base em uma única estatística da amostra; e (ii) *por intervalo*, onde se

² A *média de uma população* (μ) com N membros é $\mu = (\sum_{i=1}^N X_i) / N$, onde X denota a variável e $i = 1, 2, \dots, N$. No tocante à *média de uma amostra* ou *média de um conjunto de observações* X_i , $i = 1, 2, \dots, n$ é dada por $\bar{X} = (\sum_{i=1}^n X_i) / n$, onde \bar{X} denota a média de uma variável X considerada [Burt96].

³ O *desvio padrão* de uma amostra de observações é definida como a raiz quadrada dos desvios médios quadráticos em torno da média [Burt96], i.e. é a raiz quadrada da variância da amostra [Levi98].

considera a distribuição de amostras da média aritmética [Levi98], i.e. através do qual se constrói um intervalo que deverá, com probabilidade conhecida, conter a média populacional [Fech00].

Assim, os cálculos estatísticos permitem associar o tamanho de amostra à sua variabilidade (*desvio padrão*), a fim de gerar um intervalo de confiança (*IC*) para a média populacional. O intervalo construído a partir da estimação *por intervalo* da média da população denomina-se *intervalo de confiança* para a média. A probabilidade de que esse intervalo contenha o verdadeiro valor da média denomina-se *nível* ou *grau de confiança*, sendo representado por $(1-a)$, onde a representa a probabilidade de erro ao se afirmar que o intervalo contém o verdadeiro valor do parâmetro.

É possível calcular intervalos para qualquer grau de confiança almejado, embora intervalos de confiança de 95% sejam bastante usuais. Se o processo de amostragem da população for probabilístico, pode-se seguramente afirmar que um intervalo de confiança de 95% incluirá a média da população. Ao gerar vários intervalos de confiança de 95% a partir de várias séries de dados, pode-se, inclusive, esperar que tais intervalos incluam em 95% dos casos a média populacional verdadeira e não a inclua nos 5% restantes. Todavia, como normalmente não se conhece a média populacional, nunca é possível saber quando isto acontece [Motu99].

As estimativas *por ponto* são usualmente utilizadas nos casos em que é necessário, pelo menos aproximadamente, conhecer o valor do parâmetro de interesse para um dado propósito (e.g. uso em uma expressão analítica). Por outro lado, nos casos em que a meta final do estudo estatístico é a determinação de um dado parâmetro, adota-se geralmente a estimação *por intervalo*, tendo em vista que a probabilidade de que a estimativa *por ponto* venha a coincidir com o verdadeiro valor do parâmetro é, em geral, nula ou quase nula. Dito de outro modo, é quase certo que se esteja cometendo um erro de estimação ao se proceder a estimação por ponto de um parâmetro populacional [Fech00].

M3 Noções Conceituais de Inferência Estatística Paramétrica

Diversos problemas práticos que envolvem intervenções experimentais (e.g., medição da concentração de monóxido de carbono em diferentes áreas de um município, medição do desempenho de diferentes categorias de usuários de um aplicativo de *software*) suscitam a comparação entre amostras relativas a alguma característica de interesse, quer tais amostras provenham estatisticamente de uma única população ou de populações distintas.

Quando os resultados obtidos a partir da observação de um determinado universo amostral apresentam diferenças, usualmente se espera poder analisá-las, a fim de fazer inferências sobre um parâmetro (ou parâmetros) da população de interesse para o contexto estudado. Sabe-se que grande parte do raciocínio estatístico foi desenvolvido com propósitos direcionados para o contexto do controle de qualidade, onde qualquer análise exige questionamentos condicionais de aceitação ou rejeição, seja de um lote de componentes eletrônicos de uma indústria, de uma série de opções levantadas por uma equipe de planejamento urbano ou de um grupo de indicadores da usabilidade de um produto. O respaldo lógico utilizado para a obtenção de respostas a questionamentos dessa natureza é denominado *teste de hipóteses* [Motu99].

O teste de hipóteses parte de alguma teoria, questionamento ou afirmativa sobre o valor de um determinado parâmetro da população [Levi98] de valor desconhecido, e.g. θ , para o qual se deseja fazer inferências. Atribui-se-lhe um valor hipotético, e.g. $\theta = \theta_0$. Coleta-se uma amostra probabilística de tamanho n , que assume os valores x_1, x_2, \dots, x_n , para a qual se calcula um estimador, $\hat{\theta}$. Com base em nesse estimador, avalia-se a hipótese formulada, a fim de determinar se o valor de $\hat{\theta}$ suporta ou não a argumentação de que $\theta = \theta_0$. Se o valor do estimador estiver próximo de θ_0 , assume-se que a hipótese é verdadeira. Caso contrário, haverá duas questões a considerar: (i) a hipótese $\theta = \theta_0$ é realmente falsa; e (ii) o estimador difere de θ_0 devido a um erro de amostragem.

A Fig. 46 ilustra um esquema conceitual genérico do teste de hipóteses, enquanto o Quadro 42 sintetiza as etapas de um teste de hipóteses clássico.

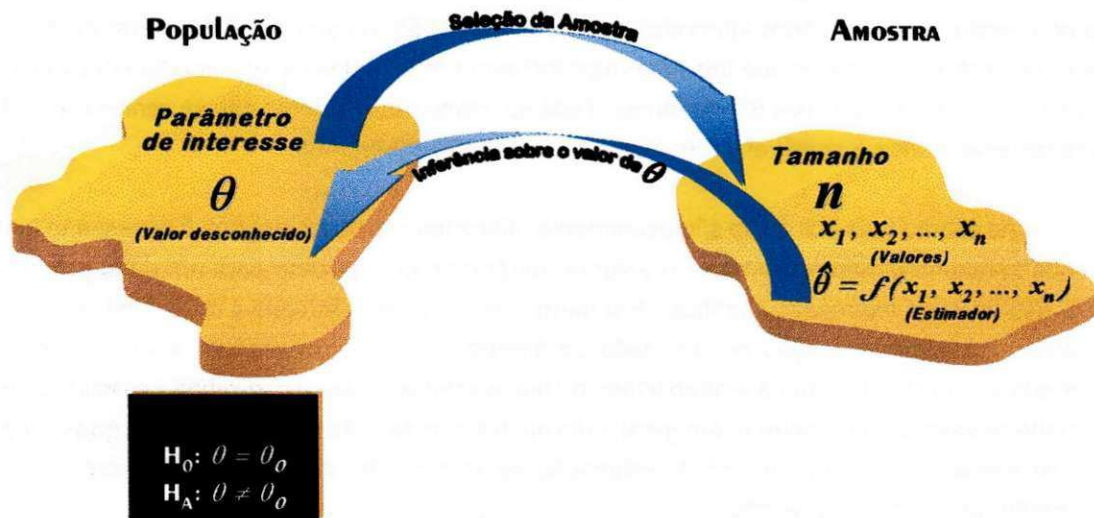


Fig. 46 – Representação Genérica de um Teste de Hipóteses Estatístico.

Quadro 42 – Discriminação das Etapas de um Teste de Hipóteses Clássico

TESTE DE HIPÓTESES CLÁSSICO	
ETAPA	ESPECIFICAÇÃO
01	Formulação de hipóteses
02	Especificação da estatística da amostra e sua distribuição amostral
03	Seleção de um nível de significância
04	Construção de uma regra de decisão
05	Computação do(s) valor(es) do teste estatístico
06	Decisão

A Fig. 46 apresenta uma síntese das declarações de um teste de hipóteses clássico, na qual a declaração que confirma a conjectura inicial ($H_0: \theta = \theta_0$) representa a *hipótese nula*, enquanto aquela que a refuta ($H_A: \theta \neq \theta_0$) representa a *hipótese alternativa*. A *hipótese nula* é sempre testada, enquanto a *hipótese alternativa*, desenvolvida como o oposto da *hipótese nula*, representa a conclusão apoiada, caso a *hipótese nula* seja rejeitada [Burt96, Levi98]. Em suma, o

teste de hipóteses é um procedimento através do qual se decide se um conjunto de dados obtidos a partir de uma amostra satisfaz ou falha como suporte a uma hipótese ou um conjunto de hipóteses de interesse (e.g., se os dados obtidos no teste usabilidade conduzido no âmbito desta pesquisa suportam as hipóteses que serão apresentadas a seguir).

Na prática, o resultado de um teste de hipóteses resulta em uma declaração de que a hipótese é suportada ou não pelos dados coletados a partir de um experimento laboratorial ou de campo. Na verdade, há alguns detalhes a considerar. Conquanto esta seja a essência do processo, universidades, empresas e, de um modo geral, a comunidade interessada nos resultados de uma pesquisa experimental não aceitam como resultados *apenas* as declarações de que uma hipótese ou um conjunto de hipóteses de interesse devem ser aceitas ou rejeitadas, à luz do processamento e análise estatísticos a que os dados do experimento foram submetidos.

Quando aplicado a situações práticas, o teste de hipóteses estatístico determina usualmente o curso de uma linha de ação [Sabo99], e.g., a otimização da interface de um aplicativo de *software* com o usuário, visando a incrementação das vendas do produto e, por conseguinte, o aumento dos lucros. Portanto, se as conclusões de um procedimento de teste de hipóteses estiverem incorretas, o curso da linha de ação nele fundamentada poderá implicar resultados potencialmente desvantajosos ou economicamente desastrosos. É necessário ter em mente que se corre o risco de formular conclusões incorretas ao utilizar estatísticas de amostras no processo de tomada de decisão sobre parâmetros da população da qual foram extraídas [Coch77, Burt96, Levi98, Troc99].

O Quadro 43 apresenta as situações efetivas e as decisões estatísticas associadas, sintetizando os resultados possíveis em um teste de hipóteses.

Quadro 43 – Matriz de Decisão de Inferência Estatística

 Decisão Estatística	 Situação Efetiva	
	H_0 é VERDADEIRA H_1 é FALSA Na realidade: ✦ <u>Não existe</u> relação ✦ <u>Não se constata</u> diferença ou ganho ✦ A teoria hipotetizada é <u>incorreta</u>	H_0 é FALSA H_1 é VERDADEIRA Na realidade: ✦ <u>Existe</u> uma relação ✦ <u>Constata-se</u> uma diferença ou ganho ✦ A teoria hipotetizada está <u>correta</u>
Se: ✦ <u>Não existe</u> relação ✦ <u>Não se constata</u> diferença ou ganho ✦ A teoria hipotetizada é <u>incorreta</u> Então: ✦ A hipótese nula deverá ser ACEITA ✦ A hipótese alternativa deverá ser REJEITADA	$1 - \alpha$ NÍVEL DE CONFIANÇA (Confidence Level) Probabilidade de se <u>aceitar</u> a hipótese <u>nula</u> quando esta é VERDADEIRA , devendo pois ser, de fato, ACEITA	β ERRO DO TIPO II (Type II Error) Probabilidade de se <u>aceitar</u> a hipótese <u>nula</u> quando esta é FALSA , devendo ser, pelo contrário, REJEITADA
Se: ✦ <u>Existe</u> relação ✦ <u>É constatada</u> uma diferença ou ganho ✦ A teoria hipotetizada está <u>correta</u> Então: ✦ A hipótese nula deverá ser REJEITADA ✦ A hipótese alternativa deverá ser ACEITA	α ERRO DO TIPO I (Type I Error) Probabilidade de se <u>rejeitar</u> a hipótese <u>nula</u> quando esta é VERDADEIRA , devendo ser, pelo contrário, ACEITA	$1 - \beta$ EFICÁCIA (Power) Probabilidade de se <u>rejeitar</u> a hipótese <u>nula</u> quando esta é FALSA , devendo pois ser, de fato, REJEITADA

Conforme se apreende do Quadro 43, assim como da leitura da sub-seção anterior, quanto maior o valor considerado para o *nível de confiança* ($1-\alpha$), mais rigorosa será a condução do teste de hipóteses. Por outro lado, quanto menor o valor considerado para o *nível de confiança* ($1-\alpha$), maior será a probabilidade de que ocorra um *erro do tipo I* (α) no resultado do teste de hipóteses. Deste modo, o valor escolhido para α definirá o *nível de significância* do teste de hipóteses, assim como seu *nível de confiança* (e.g., $\alpha=0,05$ significa que a probabilidade de que ocorra um erro do tipo I é baixa e não superior a 5%, i.e. há 95% de chances de que a *hipótese nula* não seja rejeitada quando for, de fato, verdadeira).

Adicionalmente, apreende-se que quanto maior for a probabilidade de que ocorra um *erro do tipo I* (α), menor será a probabilidade de que ocorra um *erro do tipo II* (β) e maior será a *eficácia* do teste, uma vez que se rejeitará mais freqüentemente a *hipótese nula* (aceitando-se a *hipótese alternativa*) havendo, por conseguinte, maior probabilidade de aceitação da *hipótese alternativa* quando for, de fato, verdadeira [Troc99].

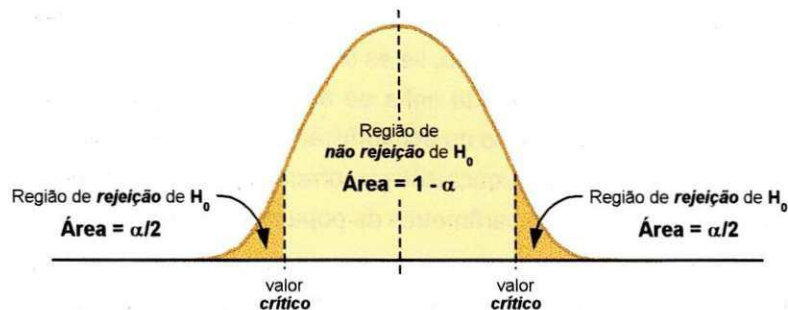


Fig. 47 – Divisão da distribuição de amostragem da estatística de um Teste de Hipóteses.

A Fig. 47 possibilita a visualização das regiões resultantes da divisão da distribuição de amostragem da estatística do teste de hipóteses. A região em que a *hipótese nula* não pode ser rejeitada é denominada *região de não rejeição*. Processos de tomada de decisão fundamentados na *hipótese nula* requerem a determinação prévia do valor correspondente à fronteira que separa as duas regiões da distribuição de amostragem da estatística do teste, denominado *valor crítico*.

Por outro lado, a região que contém valores da estatística do teste improváveis de ocorrerem se a *hipótese nula* for verdadeira é denominada *região de rejeição*, sendo também conhecida como *região crítica*. É conveniente observar que esses não serão tão improváveis de ocorrer caso a *hipótese nula* seja falsa. Deste modo, ao observar um valor da estatística do teste que se encontre nesta região, deve-se rejeitar a hipótese nula, visto que o referido valor será improvável se a *hipótese nula* for verdadeira.

M4 Teste F ANOVA: Fator Único para Diferenças nas Médias Aritméticas c

Ao realizar experimentos que envolvam categorias distintas de indivíduos ou itens, um dos principais interesses da pesquisa científica recai geralmente na confrontação dos resultados obtidos para os diferentes grupos considerados. Diversas aplicações práticas industriais ou acadêmicas envolvem estudos experimentais que consideram grupos de indivíduos ou itens relacionados a um *fator* de interesse, e.g. o desempenho de diferentes categorias de indivíduos ao interagirem com uma aplicação de *software* por ou a opinião dessas categorias de indivíduos sobre

um aspecto da interação com a aplicação de *software*.

A mensuração de um fator como o *desempenho no uso* implicará a seleção de indicadores quantitativos (e.g., tempo de execução de uma tarefa, número de ações incorretas durante a realização de uma tarefa), cada um dos quais poderá assumir diferentes valores ou níveis *numéricos* (e.g. tempos de execução iguais a 550 s, 623 s, 482 s, etc.). Por outro lado, a sondagem da opinião de diferentes categorias de indivíduos sobre um aspecto ou um conjunto de aspectos da interação com uma dada aplicação de *software* (e.g., facilidade de uso, eficácia das mensagens de erro) implicará a aquisição de opiniões associadas a escalas de níveis *categorizados* ou *semânticos* (e.g., *muito fácil, fácil, nem fácil nem difícil, difícil, muito difícil*).

Estudos experimentais dessa natureza, planejados a partir de um *fator*, nos quais os sujeitos ou unidades experimentais são atribuídos aleatoriamente a grupos de um único fator, recaem numa categoria denominada de *modelo inteiramente aleatório* ou, simplesmente, *modelo único* [Levi98], para os quais se pode empregar uma metodologia conhecida como *análise da variância* ou, de forma abreviada, ANOVA⁴ para comparar as médias aritméticas dos grupos. O teste ANOVA de *fator único* possibilita a comparação de três ou mais grupos quando os dados são categorizados segundo um único fator [Levi98, Motu99], e.g. o tempo de empacotamento de café moído, em embalagens de 250g, por três máquinas de diferentes fabricantes ou o tempo de execução de uma tarefa por usuários *principiantes, intermediários e experientes* de uma aplicação de *software*. Quando os dados são categorizados segundo dois fatores, como ilustra o exemplo anterior considerando diferenças entre usuários dos sexos masculino e feminino, a comparação é feita a partir do teste ANOVA de *dois fatores* [Coch77, Burt96, Levi98, Motu99].

Levine *et al.* [Levi98] comentam que o termo *análise da variância* pode parecer impróprio para uma metodologia cujo objetivo é analisar diferenças entre médias aritméticas, embora ressaltem que a análise da variação de séries de dados, *entre grupos (between groups)* ou *dentro de cada um dos grupos categorizados - grupos c (within groups)*, possibilita a formulação de conclusões sobre possíveis diferenças nas médias aritméticas dos grupos.

O teste ANOVA de *fator único* parte do pressuposto de que os grupos *c* ou os níveis do fator em estudo representam populações cujas medidas (i) foram retiradas aleatória e independentemente, (ii) seguem uma distribuição normal e (iii) têm variâncias equivalentes. Deste modo, a *hipótese nula* de nenhuma diferença nas médias aritméticas das populações, i.e. $H_0: \mu_1 = \mu_2 = \dots = \mu_c$, pode ser testada em relação à *hipótese alternativa* de que nem todas as médias aritméticas das populações *c* são iguais, i.e. $H_1: \text{nem todas as } \mu_j \text{ são iguais}$, onde $j = 1, 2, \dots, c$.

A *variação total* é geralmente representada pela *soma total dos quadrados (STQ)*, definida como a soma dos quadrados das diferenças entre cada valor e a média geral ou grande média, i.e.

$$\text{Variação total (STQ)} = \sum_{j=1}^c \sum_{i=1}^{n_j} (X_{ij} - \bar{X})^2 \quad (2)$$

onde $\bar{X} = \frac{\sum_{j=1}^c \sum_{i=1}^{n_j} X_{ij}}{n}$ é a *média geral* ou *grande média*, X_{ij} corresponde à *i-ésima* observação no grupo ou no nível *j*, n_j é o número de observações no grupo *j*, n é o somatório de observações de todos

⁴ Abreviação do termo inglês *ANalysis Of Variance*.

os grupos, i.e. $n = n_1 + n_2 + n_c$ e c é número de grupos ou níveis do fator de interesse [Levi98].

A *variação entre grupos*, usualmente denominada *soma dos quadrados entre grupos* (SQE), é a soma dos quadrados das diferenças entre a média aritmética da amostra de cada grupo e a média geral ou grande média, i.e.

$$\text{Variação entre grupos (SQE)} = \sum_{j=1}^c n_j (\bar{X}_j - \bar{X})^2 \quad (3)$$

onde c é número de grupos ou níveis que estão sendo comparados, n_j é o número de observações no grupo ou no nível j , \bar{X}_j é a média aritmética da amostra ou do grupo j e \bar{X} é a *média geral* ou *grande média* [Levi98].

A *variação dentro do grupo*, em geral conhecida como *soma dos quadrados dentro do grupo* (SQD), mede a diferença entre cada observação e a média aritmética de seu próprio grupo, acumulando os quadrados dessas diferenças sobre todos os grupos, sendo calculada através da equação:

$$\text{Variação dentro do grupo (SQD)} = \sum_{j=1}^c \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2 \quad (4)$$

onde X_{ij} corresponde à *i-ésima* observação no grupo ou no nível j e \bar{X}_j é a média aritmética da amostra ou do grupo j [Levi98].

Vale a pena salientar que, como os níveis c do fator estão sendo comparados, há $c - 1$ graus de liberdade associados à soma dos quadrados entre grupos. Visto que cada um dos níveis c contribui em $n_j - 1$ graus de liberdade e

$$\sum_{j=1}^c (n_j - 1) = n - c \quad (5)$$

há $n - c$ graus de liberdade associados à soma dos quadrados dentro dos grupos. Além disso, há $n - 1$ graus de liberdade associados ao total da soma dos quadrados, uma vez que cada observação X_{ij} está sendo comparada à média geral ou grande média, com base em todas as n observações.

A divisão das Eqs.(2), (3) e (4) pelos respectivos graus de liberdade resulta na obtenção de três *variâncias*⁵ ou termos *quadráticos de médias*, i.e.

$$MTQ = \frac{STQ}{n - 1} \quad (6a)$$

$$MQE = \frac{SQE}{c - 1} \quad (6b)$$

$$MQD = \frac{SQD}{n - c} \quad (6c)$$

⁵ A divisão da soma dos quadrados das diferenças por seus respectivos graus de liberdade gera, segundo a definição geral de *variância*, três variâncias. Logo, todos os termos quadráticos de médias são variâncias. Daí, o termo *análise da variância* (ANOVA).

Assim, para testar a hipótese nula ($H_0: \mu_1 = \mu_2 = \dots = \mu_c$) em relação à hipótese alternativa ($H_1: \text{nem todas as } \mu_j \text{ são iguais, onde } j = 1, 2, \dots, c$) calcula-se a estatística de teste F a partir da razão entre os termos quadráticos de médias *entre grupos* (MQE) e *dentro do grupo* (MQD), i.e.

$$F = \frac{MQE}{MQD} \quad (7)$$

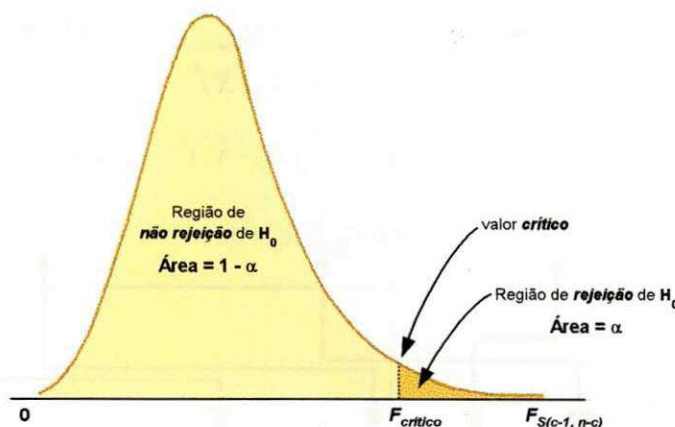


Fig. 48 – Regiões de rejeição e não rejeição para o teste de H_0 a partir do procedimento ANOVA.

Como se pode verificar observando a Fig. 48, a estatística F segue uma *distribuição F*, com $c - 1$ e $n - c$ graus de liberdade. A hipótese nula pode ser rejeitada para um determinado nível de significância (α) se a estatística do teste F exceder o valor crítico da cauda superior $F_{S(c-1, n-c)}$, de modo que a regra de decisão será:

- rejeitar H_0 , se $F > F_{S(c-1, n-c)}$,
- não rejeitar H_0 , caso contrário.

Em suma, o procedimento ANOVA gera um teste F , no qual a hipótese nula poderá ser rejeitada, em um nível de significância (α) especificado, se e somente se a estatística F for suficientemente grande a ponto de exceder $F_{S(c-1, n-c)}$, o valor crítico da cauda superior da *distribuição F*, que apresenta $c - 1$ e $n - c$ graus de liberdade.

Os resultados de um teste ANOVA são geralmente sintetizados em uma *tabela resumida ANOVA*, cujas entradas incluem as *fontes de variação* (i.e., entre grupos, dentro dos grupos e total), os *graus de liberdade*, as *somas dos quadrados*, os *quadrados médios* (i.e., as variâncias) e a *estatística F* calculada. Além dessas entradas, a maioria dos aplicativos de *software* destinados ao processamento estatístico de dados (e.g., *Statistica*, *GraphPad Prism*, *DADiSP*), assim como diversas planilhas eletrônicas (e.g., *Excel*, *QuattroPro*), incluem na *tabela resumida ANOVA* o *valor p* (ou *valor P*), que representa a probabilidade de que seja obtida uma *estatística F* suficientemente elevada ou superior àquela que se obteria se a hipótese nula fosse verdadeira.

A Fig. 49 compara duas tabelas resumidas ANOVA. A primeira delas corresponde ao arranjo recomendado por diversos autores (e.g. Levine *et al.* [Levi98], Motulski [Motu99]). A segunda foi gerada pelo *MS Excel2000* após a execução do teste ANOVA a partir dos valores associados ao número de erros repetidos, registrados ao longo da execução da **Tarefa 4** para as três categorias de participantes do teste de usabilidade conduzido no âmbito desta pesquisa (vide Tabela 2 para verificação dos dados).

FONTE DE VARIACÃO	GRAUS DE LIBERDADE	SOMA DOS QUADRADOS	QUADRADOS MÉDIOS	F
Entre grupos	$c - 1$	$SQE = \sum_{j=1}^c n_j (\bar{X}_j - \bar{X})^2$	$MQE = \frac{SQE}{c - 1}$	$F = \frac{MQE}{MQD}$
Dentro dos grupos	$n - c$	$SQD = \sum_{j=1}^c \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2$	$MQD = \frac{SQD}{n - c}$	
Total	$n - 1$	$STQ = \sum_{j=1}^c \sum_{i=1}^{n_j} (X_{ij} - \bar{X})^2$		

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	172,5375	2	86,2688	7,054993784	0,002537319	3,251926728
Within Groups	452,4375	37	12,2280			
Total	624,9750	39				

Fig. 49 – Tabelas resumidas ANOVA.

Observando as tabelas acima ilustradas, verifica-se que a tabela gerada pelo *MS Excel* apresenta, além das informações indicadas na tabela superior, o valor *P* e o valor crítico de *F* (vide Fig. 48). As informações contidas na segunda e terceira colunas da tabela superior aparecem invertidas na tabela inferior, conforme indicado pelas setas. A partir das informações contidas nas três últimas colunas da tabela inferior, que aparecem destacadas sobre fundos amarelo (F e F_{crit}) e laranja (*P-value*), decide-se se a hipótese nula será ou não rejeitada. Se $F > F_{crit}$ ou, alternativamente, se $P\text{-value} < \alpha$, então H_0 poderá ser rejeitada, como ocorre neste caso ($\alpha = 0,05$).