

Universidade Federal de Campina Grande  
Centro de Engenharia Elétrica e Informática  
Coordenação de Pós-Graduação em Ciência da Computação

Avaliação da Utilização de *Leaderboards* em  
Disciplinas Introdutórias de Programação

Kláudio Henrique Mascarenhas Medeiros

Dissertação submetida à Coordenação do Curso de Pós-Graduação em  
Ciência da Computação da Universidade Federal de Campina Grande -  
Campus I como parte dos requisitos necessários para obtenção do grau  
de Mestre em Ciência da Computação.

Área de Concentração: Ciência da Computação

Linha de Pesquisa: Engenharia de Software

Dalton Dario Serey Guerrero e Jorge Cesar Abrantes de Figueiredo  
(Orientadores)

Campina Grande, Paraíba, Brasil

©Kláudio Henrique Mascarenhas Medeiros, 01/11/2016

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL DA UFCG

M488a Medeiros, Kláudio Henrique Mascarenhas.  
Avaliação da utilização de *Leaderboards* em disciplinas introdutórias de programação / Kláudio Henrique Mascarenhas Medeiros. – Campina Grande, 2016.  
67 f. : il.

Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e informática, 2016.

"Orientação: Prof. Dr. Dalton Dario Serey Guerrero, Prof. Dr. Jorge Cesar Abrantes de Figueiredo.

Referências.

1. Ensino de Programação. 2. Gamificação. 3. *Leaderboard*. I. Guerrero, Dalton Dario Serey. II. Figueiredo, Jorge Cesar Abrantes de. III. Título.

CDU 004.4(07)(043)

## Resumo

Dentre as práticas importantes para alunos de programação introdutória está a realização de exercícios. Existem disciplinas de programação que disponibilizam aos alunos um grande conjunto de problemas. Apesar dessa disponibilidade, existe o problema de que as disciplinas contêm alunos que não são motivados a resolvê-los. A gamificação consiste no uso de elementos e dinâmicas existentes em jogos em contextos que não são jogos e nos últimos anos tem sido considerada uma possível solução para o problema de desengajamento de alunos em contextos educacionais. Um dos elementos de jogos utilizados em contextos gamificados é o *leaderboard*. Os *leaderboards* são tabelas que exibem indivíduos ordenados de acordo com um critério de pontuação e criam competições que podem estimular os alunos a serem mais ativos. Neste trabalho, nós propomos e avaliamos um design de *leaderboards* para disciplinas introdutórias de programação, com o intuito de incentivar os alunos a resolverem mais exercícios de programação. Nós elicitamos requisitos baseados nos aspectos positivos e negativos dos *leaderboards* e tomamos decisões de design para atendê-los. A avaliação do design foi feita através de um experimento de curta duração e um estudo longitudinal durante um período letivo. Nós encontramos indícios de que os *leaderboards* exercem efeito positivo sobre os alunos, fazendo-os exercitar mais. A atribuição de pontuação igual para todos os tipos de problemas permite que mais alunos tenham chances de alcançar altas posições nos *leaderboards*. A implantação de *leaderboards* de menor duração incentivou alunos menos ativos na disciplina, mesmo que por poucos dias.

## **Abstract**

Among the important practices to students of introductory programming is the exercise solving. There are programming courses that offer a large set of problems to students. Despite this availability, there are cases of courses that have students that are not motivated to solve these problems. Gamification is the use of game elements and dynamics in non-game contexts and recently has been considered to be a possible solution for the problem of disengagement of students on educational contexts. One of the game elements used in gamified contexts is the leaderboard. The leaderboards are charts that shows individuals ranked accordingly to a score criteria and creates competition environments that may stimulate students to be more active. In this work, we propose and evaluate a leaderboard design for introductory programming courses, in order to motivate the students to solve more programming exercises. We elicited requisites based on the positive and negative aspects of leaderboards and made design decisions for fulfill those requirements. The evaluation of the design is done through an short-length experiment and a longitudinal study during a semester. We found evidences that the leaderboards have a positive effect on the students, motivating them to do more exercises. The attribution of the same score for all exercises allowed more students to reach high ranks on the leaderboards. The implantation of short-length leaderboards motivated less-active students on the course, even for a few days.

## Agradecimentos

Agradeço primeiramente aos meus pais, Kalina e Laerte, por todo o amor e dedicação que me ofereceram. São excelentes pais e espero seguir seus exemplos, hoje e sempre.

Agradeço meus orientadores, Dalton e Jorge, pelos ensinamentos, dicas e pelos momentos no mestrado e em sala de aula.

Agradeço ao curso de pós-graduação em Ciência da Computação da UFCG pelo acolhimento, pelo suporte e por todo o aprendizado que tive nas disciplinas que cursei.

Agradeço a Mariana, minha namorada, pelo amor, pela companhia, pelo apoio e pelo carinho que são tão importantes para mim.

Agradeço aos meus colegas do SPLab, em especial ao grupo de pesquisa de Educação e Computação, a Alysson, Catharine, Izabela e Matheus, pelas discussões, pelos *feedbacks*, pelos desabafos e pela amizade que desenvolvemos.

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Contextualização . . . . .	1
1.2	Objetivo deste trabalho . . . . .	3
1.3	Design da pesquisa . . . . .	4
1.4	Resultados . . . . .	4
1.5	Estrutura da dissertação . . . . .	6
<b>2</b>	<b>Fundamentação teórica</b>	<b>7</b>
2.1	Jogos ( <i>games</i> ) . . . . .	7
2.2	Jogos e educação . . . . .	8
2.3	Gamificação . . . . .	9
2.4	<i>Leaderboards</i> . . . . .	10
2.5	Trabalhos relacionados . . . . .	11
<b>3</b>	<b>Design para <i>leaderboards</i></b>	<b>14</b>
3.1	Requisitos . . . . .	14
3.2	Decisões de design . . . . .	16
3.3	Avaliação do design . . . . .	19
<b>4</b>	<b>Avaliação do efeito de um <i>leaderboard</i> na quantidade de atividade desempenhada por alunos em uma aula de programação</b>	<b>20</b>
4.1	Perguntas de pesquisa . . . . .	21
4.2	Metodologia . . . . .	23
4.2.1	Participantes . . . . .	23
4.2.2	Questões . . . . .	24

4.2.3	<i>Leaderboard</i> . . . . .	24
4.2.4	Condução do experimento . . . . .	25
4.3	Resultados . . . . .	26
4.3.1	Análise das métricas de atividade . . . . .	26
4.3.2	Tentativas de problemas por unidade . . . . .	31
4.3.3	Posição dos alunos . . . . .	35
4.4	Discussões . . . . .	36
4.4.1	Análise das métricas de atividade . . . . .	36
4.4.2	Tentativas de problemas por unidade . . . . .	37
4.4.3	Posição dos alunos . . . . .	38
4.5	Ameaças à validade . . . . .	38
<b>5</b>	<b>Avaliação da utilização de <i>leaderboards</i> durante um período letivo</b>	<b>40</b>
5.1	Perguntas de pesquisa . . . . .	40
5.2	Metodologia . . . . .	42
5.2.1	Participantes . . . . .	42
5.2.2	Métricas . . . . .	42
5.2.3	Design dos <i>leaderboards</i> . . . . .	43
5.3	Resultados . . . . .	44
5.3.1	Visualização dos <i>leaderboards</i> por tipo . . . . .	44
5.3.2	Visualização dos <i>leaderboards</i> e resolução de exercícios no tempo .	47
5.3.3	Relações entre visualização de <i>leaderboards</i> e resolução de exercícios	49
5.4	Discussões . . . . .	54
5.4.1	Visualização dos <i>leaderboards</i> por tipo . . . . .	55
5.4.2	Visualização dos <i>leaderboards</i> e resolução de exercícios no tempo .	56
5.4.3	Relação entre visualização de <i>leaderboards</i> e resolução de exercícios	57
<b>6</b>	<b>Conclusões e trabalhos futuros</b>	<b>59</b>
6.1	Conclusões . . . . .	59
6.2	Trabalhos futuros . . . . .	60

# Lista de Figuras

4.1	<i>Boxplot</i> da métrica <i>READS</i> entre os dois grupos . . . . .	27
4.2	<i>Boxplot</i> da métrica <i>SUBS</i> entre os dois grupos . . . . .	27
4.3	<i>Boxplot</i> da métrica <i>PROBLEMS</i> entre os dois grupos . . . . .	28
4.4	<i>Boxplot</i> da métrica <i>SOLVED</i> entre os dois grupos . . . . .	28
4.5	<i>Boxplot</i> da métrica <i>TIME</i> entre os dois grupos . . . . .	29
5.1	Quantidade de visualizações do <i>leaderboard</i> geral feitas pelos alunos . . . .	45
5.2	Quantidade de visualizações do <i>leaderboard</i> semanal feitas pelos alunos . .	46
5.3	Quantidade de visualizações do <i>leaderboard</i> diário feitas pelos alunos . . .	46
5.4	Quantidade de visualizações dos <i>leaderboards</i> feitas por dia pelos alunos .	47
5.5	Quantidade de exercícios resolvidos pelos alunos em todos os dias . . . . .	48

# Lista de Tabelas

4.1	Médias de cada métrica para cada grupo . . . . .	29
4.2	Desvios padrão de cada métrica para cada grupo . . . . .	30
4.3	p-valores dos testes de hipótese para cada métrica entre os grupos . . . . .	31
4.4	Somatório de problemas abordados por unidade por alunos . . . . .	32
4.5	p-valores dos testes de hipótese de problemas abordados por unidade . . . . .	32
4.6	Unidades dos problemas tentados por alunos do grupo experimental . . . . .	33
4.7	Unidades dos problemas tentados por alunos do grupo de controle . . . . .	34
4.8	Resultado final do <i>leaderboard</i> . . . . .	35
5.1	Quantidade de visualizações dos <i>leaderboards</i> separadas por tipo . . . . .	44
5.2	Correlação entre visualização de <i>leaderboard</i> e resolução de exercício dos alunos . . . . .	50
5.3	Dados sobre visualização de <i>leaderboards</i> e resolução de exercícios dos alunos	50

# Capítulo 1

## Introdução

### 1.1 Contextualização

A realização de exercícios representa uma prática importante para o aluno em disciplinas de programação introdutória [19]. Em geral, exercícios de programação correspondem à pequenos problemas que cobram dos alunos a construção de códigos pequenos que solucionem esses problemas [46]. Através desses exercícios, o aluno pode aplicar e fixar conceitos, planejar e implementar estratégias de resolução de problemas e, assim, desenvolver uma melhor preparação para programação. A quantidade de exercícios resolvidos pelo aluno apresenta uma forte correlação com o seu sucesso [11].

Nessas disciplinas, é possível perceber alunos com diferentes motivações diante dos exercícios. Existem alunos que são bastante engajados com a resolução de problemas por conta da importância dos exercícios para o seu aprendizado ou, simplesmente, por interesse ou vontade própria de programar. Por outro lado, existem alunos que não possuem tanta motivação e realizam exercícios com pouca frequência.

Uma decisão possível e comum para estimular a realização de exercícios por parte do estudante é a de determinar que parte de sua nota final seja proveniente dos exercícios realizados por ele durante a disciplina. Essa decisão consegue fazer com que os alunos exercitem, mas os estudantes podem se sentir pressionados por conta da influência dos exercícios em sua nota final. Deste modo, a realização de exercícios se torna uma atividade praticamente mandatória, existindo uma punição caso ele não os resolva. Diante dessa situação, é ideal considerar a utilização de abordagens mais lúdicas para incentivar a realização de exercícios.

Dentro das abordagens lúdicas existentes, os jogos representam uma das formas de entretenimento mais interessantes e engajadoras. Os jogos disponibilizam aos jogadores mundos, personagens, histórias e missões, com os quais os jogadores se envolvem e demonstram o empenho e a dedicação que, muitas vezes, não demonstram em situações do cotidiano. Com o reconhecimento do potencial dos jogos e o crescimento de sua popularidade, os acadêmicos têm tido interesse nesse tipo de entretenimento, buscando compreender quais características os tornam motivantes e como jogos podem ser concebidos e utilizados no processo de ensino e aprendizagem. Esses esforços deram origem a jogos como os jogos sérios, que prezam pelo desenvolvimento de jogos com objetivos educacionais como aprendizagem de conceitos e treinamento. Nos últimos anos, cresceu na indústria e na academia uma alternativa mais simples e menos custosa ao desenvolvimento de jogos completos com fins educativos, que recebeu o nome de gamificação.

De acordo com Deterding et al.[15], gamificação <sup>1</sup> é definida como sendo o uso de elementos de *game design* (e.g. mecânicas, dinâmicas e estéticas [27]) em contextos que não são jogos. É uma estratégia que tem objetivos como os de motivar e engajar usuários, promover aprendizagem e resolução de problemas [31]. Em contextos educacionais, a gamificação tem sido implantada com o intuito de melhorar o processo de ensino e aprendizagem, tornando-o mais cativante e eficiente, e para motivar estudantes a diversas atividades, dentre as quais a realização de exercícios [7].

Dentre os diferentes elementos e dinâmicas presentes em jogos, o *leaderboard* é um dos mais utilizados em contextos gamificados [25]. Os *leaderboards* são tabelas que exibem os participantes de um contexto ordenados de acordo com um determinado critério. Eles permitem que cada indivíduo possa se comparar com os demais, possibilitando-o inferir seu desempenho em relação aos seus pares. A implantação de *leaderboards* dá origem a dinâmicas competitivas em contextos educacionais, que podem facilitar a visualização do progresso e aumentar a motivação e interação social dos alunos. Entretanto, a utilização de *leaderboards* exige cuidados, pois podem exercer o efeito contrário ao desejado, desmotivando alunos e fazendo-os vivenciar situações de pressão e desconforto devido à exposição de informação.

---

<sup>1</sup>No inglês, *gamification*.

## 1.2 Objetivo deste trabalho

O objetivo deste trabalho é propor e avaliar um design de *leaderboards* para disciplinas introdutórias de programação, com o intuito de motivar estudantes a resolverem mais exercícios de programação. O design proposto busca explorar e ampliar os pontos positivos de *leaderboards* e dinâmicas competitivas (e.g. chances de vitória) e amenizar os pontos negativos (e.g. exposição de informação dos participantes para seus pares).

Para atingirmos nosso objetivo, desenvolvemos *leaderboards* e os implantamos em um sistema de correção automática utilizado por alunos da disciplina de Laboratório de Programação I, pertencente à grade curricular do curso de Ciência da Computação da Universidade Federal de Campina Grande (UFCG). Nesse sistema, os alunos podem acessar exercícios e submeter soluções para eles, recebendo do sistema um *feedback* dessa solução. Nós usamos os dados provenientes dos registros de interação dos alunos com esse sistema para avaliação do design.

Este trabalho busca responder as seguintes perguntas de pesquisa:

- **Pergunta de pesquisa 1:** A existência de *leaderboards* tem efeito na quantidade de atividade realizada pelos alunos em disciplinas introdutórias de programação?
- **Pergunta de pesquisa 2:** Quais os impactos da utilização de um critério de pontuação que pontua todos os exercícios igualmente?
- **Pergunta de pesquisa 3:** A implantação de *leaderboards* de curta duração incentiva os alunos a exercitarem?
- **Pergunta de pesquisa 4:** Qual o nível de interesse dos alunos em relação aos *leaderboards* no decorrer de um período letivo?
- **Pergunta de pesquisa 5:** Qual a relação entre visualização dos *leaderboards* e resolução de exercícios por parte dos alunos?

Para as perguntas 1 e 2, fizemos um estudo experimental durante uma aula prática de programação introdutória. Para as perguntas 3, 4 e 5, realizamos um experimento longitudinal durante uma disciplina de programação introdutória.

## 1.3 Design da pesquisa

Nós elaboramos e avaliamos o design de *leaderboards* através de dois estudos. No primeiro estudo, realizamos um experimento controlado com alunos matriculados em uma turma de programação introdutória. Buscamos nesse experimento observar se os *leaderboards* incentivam os alunos a realizarem mais exercícios de programação. Além disso, avaliamos os impactos de um critério de pontuação que considera que todos os exercícios possuem o mesmo valor, como uma possível maneira de amenizar as disparidades entre alunos com diferentes níveis de domínio do assunto na competição. Nós dividimos os alunos em dois grupos, sendo um grupo experimental com acesso a um *leaderboard* projetado na parede da sala contendo a classificação dos alunos, e um grupo de controle que não tinha conhecimento sobre a existência do *leaderboard*.

Em um segundo momento, realizamos um estudo longitudinal durante um período letivo em uma disciplina introdutória de programação, aumentando o tempo de contato dos alunos com os *leaderboards* para 70 dias e adotando mais critérios e decisões de design. Nós implementamos e disponibilizamos três tipos de *leaderboard* distintos, disponíveis em links acessíveis pelos alunos, que contabilizaram os exercícios resolvidos por eles em diferentes intervalos de tempo. Além disso, modificamos os *leaderboards* com o objetivo de restringirmos a visibilidade dos alunos. Essa modificação estabeleceu que cada aluno só visualizaria os alunos empatados com ele, em posições vizinhas à dele e os alunos nas posições do topo. Neste estudo, contabilizamos a frequência de visualizações recebidas pelos *leaderboards* e resolução de exercícios no tempo, além da quantidade de visualizações feitas por cada aluno à cada tipo de *leaderboard* e as relações das visualizações com os exercícios resolvidos pelos alunos.

## 1.4 Resultados

Os nossos estudos mostram que a utilização de *leaderboards* incentiva os alunos a fazerem mais exercícios de programação, além de que decisões de design propostas potencializam os benefícios dos *leaderboards*. Vimos que alunos com acesso a um *leaderboard* durante uma aula submeteram mais códigos para mais problemas distintos, conseguindo também

resolver mais exercícios que um grupo sem o *leaderboard*. Nós encontramos indícios de que o critério de pontuação que pontua igualmente todos os exercícios incentiva a realização de exercícios de assuntos mais simples, sem que aconteça redução na realização de exercícios mais avançados. Nós percebemos que esse critério de pontuação permite que alunos mais atrasados tenham chances de obter altas posições no *leaderboard*, fazendo com que mais alunos tenham chances de vitória, gerando uma competição mais equilibrada.

Através da investigação dos *leaderboards* durante um período letivo, vimos que a quantidade de visualizações realizadas pelos alunos é alta no início e, conforme o tempo passa, essa quantidade vai sendo reduzida. Este fenômeno pode ter ocorrido devido ao efeito da novidade dos *leaderboards*, de modo que, conforme o *leaderboard* se torna menos novidade para o aluno, o interesse do estudante é reduzido. Além disso, as visualizações foram afetadas por fatores externos como avaliações de outras disciplinas que, principalmente no final do período letivo, fazem com que os alunos exercitem menos. Por fim, fatores internos relacionados com a estrutura da própria disciplina na qual o estudo foi conduzido, que permite os alunos seguirem ritmos de progresso distintos entre si e cumpram com a ementa da disciplina antes do tempo, também influenciaram os resultados.

Quanto às visualizações feitas a cada tipo de *leaderboard*, vimos que o *leaderboard* geral foi o mais acessado pelos alunos que resolveram mais exercícios durante o semestre, o que indica que esse tipo é o mais interessante para os alunos mais ativos. Por sua vez, os *leaderboards* semanais e diários foram os tipos mais acessados por parte dos demais alunos, que apresentaram aumentos esporádicos na quantidade diária de exercícios resolvidos e de visualizações dos *leaderboards*, realizando estas duas ações de forma alternada. Esses resultados sugerem que os *leaderboards* de menor duração incentivam mais alunos a resolverem exercícios, mesmo que durante um período de tempo curto.

Nós vimos também alunos que aumentaram significativamente o número de exercícios resolvidos após a implantação dos *leaderboards*. Uma parcela desses estudantes apresentou uma correlação alta entre visualização de *leaderboards* e resolução de exercícios, o que representa um indício que os *leaderboards* foram interessantes para esses alunos e os motivaram a exercitar.

## 1.5 Estrutura da dissertação

O capítulo 2 contém a fundamentação teórica deste trabalho. Nesse capítulo, abordamos o estado-da-arte de gamificação, com foco em trabalhos que utilizaram *leaderboards* em contextos educacionais.

No capítulo 3 propomos e abordamos um design para *leaderboards*, levantando requisitos baseados em aspectos relevantes de competições e elaborando decisões de design para atender a estes requisitos.

Nos capítulos 4 e 5 detalhamos os resultados e discussões acerca das pesquisas realizadas em relação à utilização e o impacto de decisões de design para *leaderboards* na realização de exercícios.

Por fim, o capítulo 6 traz as conclusões obtidas com este trabalho, além de perspectivas de trabalhos futuros.

# Capítulo 2

## Fundamentação teórica

Neste capítulo, discutiremos os conceitos importantes para o entendimento do presente trabalho. Dentre os conceitos, iremos abordar a gamificação, mostrando como ela surgiu, como evoluiu com o passar dos anos, quais resultados tem conseguido e quais desafios e limitações ainda possui. Dentro dos elementos de jogos presentes em gamificação, iremos estabelecer foco nos *leaderboards* e em conceitos relacionados, como competições.

### 2.1 Jogos (*games*)

Nos últimos anos, o contexto de jogos digitais passou por uma constante ascensão como mercado e como campo de pesquisa [49]. Disponíveis em consoles, celulares, computadores e em outras formas de *hardware*, os jogos digitais estão presentes no cotidiano das pessoas na forma de jogos simples para passar o tempo como *Angry Birds*<sup>1</sup> até ambientes online com milhões de jogadores jogando simultaneamente como *World of Warcraft*<sup>2</sup>. Hoje em dia, os jogos são usufruídos por pessoas de todas as idades e de ambos os sexos [3].

A literatura apresenta uma diversidade de definições para o que é um jogo. Para Salen e Zimmerman [47], um jogo é um sistema no qual jogadores estão engajados em um conflito artificial, definido por regras, que resulta em um resultado quantificável. A definição de Avedon e Sutton-Smith [4] considera os jogos como sendo atividades voluntárias regidas por regras, com um conflito entre partes e um resultado desigual entre estas partes. A definição

---

<sup>1</sup><https://www.angrybirds.com/>

<sup>2</sup><https://worldofwarcraft.com/pt-br/>

de Schell [48] aponta que um jogo é uma atividade de resolução de problemas com uma atitude lúdica. Na definição de Juul [30], os jogos são formados por regras, resultados quantificáveis, valor positivo ou negativo atribuído a cada resultado, esforço do jogador, ligação entre jogador-resultados e consequências negociáveis, que podem ou não ser refletidas no mundo real. Para McGonigal [41], um jogo é caracterizado por objetivos, regras, existência de um sistema de *feedback* e participação voluntária.

## 2.2 Jogos e educação

Dentro dos jogos, os jogadores estão engajados e se divertem com desafios longos e difíceis. Essa postura nem sempre é observada em um aluno dentro de um contexto educacional. Diante do potencial existente nos jogos de cativar, divertir e motivar pessoas, surgiu dentre os pesquisadores em educação um grande interesse na compreensão, concepção e utilização de jogos com o objetivo de melhorar o processo de ensino e aprendizagem.

Malone [40] elaborou heurísticas para a criação de interfaces intrinsecamente interessantes. No *framework* proposto por ele, existem três tipos de motivação que devem ser incentivados nos alunos: desafios, curiosidade e fantasia. Gee [24] estudou e elaborou princípios de ensino e aprendizagem que jogos possuem. Dentre esses princípios estão a interação e o *feedback* imediato, produção de conteúdo, customização, ciclos de *expertise*, informação sob demanda, colaboração e baixo impacto negativo das falhas, que são utilizadas pelo jogador como uma forma de aprender padrões e avançar no jogo. Em seu livro *Reality is Broken* [41], Jane McGonigal afirma que os jogos satisfazem vontades humanas como as de experiências de sucesso, conexão social e trabalho satisfatório e desenvolvem características como persistência, criatividade e resiliência.

A combinação de jogos e educação pode se manifestar de diferentes formas. Uma forma muito usual é a de jogo sério<sup>3</sup>, que é definido por Ritterfeld, Cody e Vorderer [44] como sendo "qualquer forma de software de jogo interativo baseado em computador para um ou mais jogadores para ser usado em qualquer plataforma e desenvolvido para promover mais que entretenimento para os jogadores". Os jogos sérios são utilizados para fins como treinamento e conscientização em contextos como os de educação, engenharia, política e plane-

---

<sup>3</sup>No inglês, *serious game*.

jamento [42]. A utilização de jogos sérios em contextos educacionais trouxe benefícios em relação ao aprendizado, motivação e melhora de habilidade dos alunos [10].

A criação de jogos completos com fins educacionais demanda muito tempo, recursos, tecnologias e aplicação de *game design*, tornando-os produtos custosos e complexos [31]. Diante desse problema, começou a se considerar uma alternativa mais simples e barata, porém ainda eficiente, aos jogos completos [14] [42], chamada dentro da literatura e na indústria de gamificação.

## 2.3 Gamificação

Segundo Deterding et al. [15], gamificação consiste na utilização de elementos de *game design* em contextos que não são jogos. Embora a primeira citação do termo “gamificação” tenha ocorrido em 2008, gamificação teve um crescimento expressivo por volta de 2010 [15]. Inicialmente, gamificação era utilizada principalmente para promover serviços e produtos nos contextos de negócios e marketing para promover produtividade e lealdade [1]. Ao mesmo tempo em que gamificação crescia na indústria e na academia, começava a se discutir sobre o seu potencial na educação [37] [42]. Em poucos anos, educação se tornou um dos principais setores de aplicação de gamificação [25] [49].

De acordo com Dicheva e Dichev [16], gamificação em educação corresponde à introdução de elementos de jogos e experiências de jogo no design de processos de aprendizagem. A adoção de gamificação em educação visa dar suporte ao ensino, além de incentivar participação, colaboração, realização de exercícios, exploração e criatividade [7], engajando os estudantes e modificando seu comportamento. As aplicações de gamificação em educação são diversas, incluindo MOOCs, sites de *e-learning* e cursos com aprendizagem híbrida [16]. Um dos maiores casos existentes de gamificação em educação é o Quest to Learn (Q2L), que corresponde a um projeto curricular que tem como objetivo a gamificação de todo o currículo escolar [2].

A revisão da literatura de Hamari, Koivisto e Sarsa [25] aponta que a utilização de gamificação traz efeitos positivos, mas que esses resultados são fortemente dependentes do contexto onde gamificação está sendo implementada e de características dos usuários que estão nesses contextos. Além disso, essa mesma revisão indica que os efeitos da gamificação

não são efeitos de longo prazo e que podem existir devido a um efeito da novidade inerente à implantação de gamificação. A literatura ainda carece de trabalhos que investigam as relações entre as motivações intrínseca e extrínseca e contextos gamificados [17] [49]. Além disso, segundo Dicheva e Dichev [16], a crescente quantidade de trabalhos inconclusivos e com resultados negativos acerca da gamificação sugerem que a técnica precisa ser trabalhada de modo mais crítico.

## 2.4 Leaderboards

O *leaderboard* é um elemento de jogo que consiste em uma tabela que exhibe indivíduos ordenados de acordo com determinados critérios e é um dos elementos de jogos mais utilizados em contextos gamificados [25]. Em muitos desses contextos, é usado em conjunto com pontos e *badges*, formando o que é conhecido na literatura como PBL (*Points, Badges e Leaderboards*). Dentro de contextos educacionais, permitem que os alunos comparem seu desempenho com o desempenho de seus colegas [9].

A teoria da comparação social [20] afirma que indivíduos avaliam suas habilidades e seus valores através da realização de comparações com outros indivíduos. As comparações realizadas através dos *leaderboards* tipicamente implicam em competições entre indivíduos [22]. A literatura de competição em educação é dividida entre defensores e opositores. Kohn afirma que competições causam ansiedade nos estudantes [34] e possuem efeito negativo nas relações humanas, auto-estima e produtividade [33]. A revisão da literatura de Johnson et al. [29] afirma que competição é menos eficiente do que abordagens cooperativas. Segundo Fülöp [21], competição pode ser benéfica se for vista de forma positiva e direcionada à motivação e desenvolvimento pelos alunos. Tauer e Harackiewicz [50] afirmam que competição pode ter efeitos positivos e que a combinação de competição e cooperação tem efeito positivo sobre a motivação intrínseca e a performance.

Existem trabalhos que categorizam os jogadores de acordo com suas preferências e motivações dentro dos jogos, e a competição é uma das mais presentes. Na taxonomia de tipos de jogador de Bartle [5], existe um perfil que é motivado por competições e por vencer seus pares chamado de *killer*. No estudo de Yee [51] foi identificado, dentre diversas atividades, o desejo de desafiar e competir com outros jogadores. O jogador com o perfil de competidor

é incentivado por um elemento de jogo como o *leaderboard*, que o permite comparar seu desempenho e se impor sobre os outros jogadores [28].

A teoria da autodeterminação <sup>4</sup> contém anos de pesquisa empírica que apontam que as recompensas e a competição incentivadas pela gamificação podem ter impacto negativo na motivação intrínseca do indivíduo [12]. A motivação intrínseca por uma atividade corresponde à motivação que é proveniente da própria atividade, ao invés de consequências externas [13]. Por sua vez, as recompensas e os *leaderboards*, em geral, são fontes de motivação extrínseca. Os efeitos dos elementos de jogos na motivação intrínseca do indivíduo ainda não estão estabelecidos na literatura [17], mas constantemente os trabalhos alertam sobre cuidados que devem ser tomados na utilização de incentivos extrínsecos, pois estes podem reduzir ou até mesmo substituir a motivação intrínseca do indivíduo [38].

Os *leaderboards* podem ter efeito informacional ou controlador no indivíduo e, de acordo com a teoria de avaliação cognitiva <sup>5</sup>, o efeito informacional está relacionado de forma positiva com a motivação intrínseca do indivíduo, enquanto que o efeito controlador está relacionado de forma negativa [45]. Se o indivíduo enxergar o *leaderboard* de forma informacional, irá se sentir mais autônomo e competente e, portanto, terá sua motivação intrínseca aumentada. Porém, se o indivíduo enxergar o *leaderboard* de forma controladora, irá sentir menos autonomia e, conseqüentemente, haverá uma redução na sua motivação intrínseca. Reeve e Deci afirmam que vencer na competição aumenta a motivação intrínseca do indivíduo, mas que aspectos controladores do contexto podem pressionar o indivíduo e reduzir a sua motivação intrínseca [43].

## 2.5 Trabalhos relacionados

A maioria dos trabalhos que utilizam *leaderboards* o fazem em conjunto com outros elementos de jogos [17] [25]. Em geral, os trabalhos utilizam vários elementos de jogos simultaneamente e discutem sobre gamificação como um todo. Esta seção irá trazer trabalhos que utilizaram *leaderboards*, de forma isolada ou em conjunto com outros elementos em contextos educacionais e que trazem resultados e discussões acerca dos *leaderboards*.

---

<sup>4</sup>No inglês, *self-determination theory*.

<sup>5</sup>No inglês, *cognitive evaluation theory*.

- Domínguez et al. [18] utilizaram *leaderboards* em um *plugin* gamificado para uma plataforma de *e-learning* para comparar os alunos em relação a quantidade de *badges* obtidos. Foi realizado um experimento que envolveu um grupo de controle e um experimental, este último com acesso ao *plugin* gamificado. Os resultados mostram que o grupo experimental obteve notas melhores em atividades práticas, mas a performance em atividades escritas e participação em sala de aula foram piores. A análise qualitativa revelou opiniões divergentes sobre os *leaderboards*. Existiram alunos que se sentiram motivados pelos *leaderboards* e fizeram tarefas por causa dele e existiram alunos que não gostaram dos *leaderboards* e do clima competitivo existente.
- Cheong, Cheong e Filippou [8] desenvolveram um quiz gamificado que continha *leaderboards*. Alunos de turmas de TI utilizaram esse quiz gamificado e responderam um *survey*, de onde foram provenientes os dados utilizados no estudo. A maioria dos estudantes afirmou que esteve engajada e feliz enquanto usava o quiz e que o uso do quiz causou um aumento na eficiência do aprendizado. Quando aos *leaderboards*, alguns alunos afirmaram que a competição foi motivante, assim como houveram alunos que se sentiram envergonhados por estarem nas últimas posições.
- Christy e Fox [9] utilizaram *leaderboards* em um teste virtual de Matemática com 80 estudantes do sexo feminino que foram divididas em grupos e expostas a diferentes *leaderboards*: um onde os homens estavam nas melhores posições, outro onde mulheres estavam nas melhores posições e um terceiro grupo sem *leaderboards*. Esse estudo buscou verificar o efeito de possíveis ameaças de estereótipo e comparações sociais na performance das estudantes. Os resultados deste estudo mostram que as mulheres do grupo com *leaderboard* liderado por mulheres tiveram um desempenho pior, mas apresentaram identificação com Matemática maior que os demais grupos.
- Hanus e Fox [26] realizaram um estudo longitudinal com alunos de dois cursos de 16 semanas, onde uma turma recebeu um currículo gamificado, e mediu satisfação, motivação e esforço dos alunos. Foi observado que alunos da turma gamificada demonstraram menos motivação e satisfação ao longo do tempo, além de notas finais piores que o grupo sem gamificação.

- Landers e Landers [36] realizaram um experimento onde alunos realizaram um projeto online e foram separados em dois grupos: um com uma versão gamificada do projeto, que incluiu *leaderboards*, e outro grupo sem *leaderboards*. Os alunos com *leaderboards* interagiram mais com o projeto e tiveram melhor performance.

Os trabalhos supracitados utilizam *leaderboards* e discutem sobre os impactos de sua utilização em contextos educacionais. O presente trabalho busca estender essa discussão através da proposição e avaliação de um design de *leaderboards* para disciplinas de programação introdutória. Para elaborarmos o design, nós levantamos aspectos positivos e negativos dos *leaderboards* e, a partir deles, elaboramos requisitos. Após levantarmos os requisitos, trabalhamos em decisões de design para atender esses requisitos. A descrição dos requisitos e das decisões de design elaboradas para atendê-los está descrito no próximo capítulo.

# Capítulo 3

## Design para *leaderboards*

Este trabalho tem como objetivo propor e avaliar um design de *leaderboards* para disciplinas introdutórias de programação, com o intuito de incentivar os alunos a resolverem mais exercícios de programação. Os *leaderboards* são tabelas que exibem os participantes de um determinado contexto ordenados de acordo com um critério de pontuação.

Os *leaderboards* possuem características positivas e negativas, como estabelecimento de competições que engajam os participantes e exposição de informações a respeito do desempenho dos participantes de modo a desmotivá-los. Baseado nessas características, nós elaboramos requisitos. Esses requisitos conduziram decisões de design tomadas para realçar e aumentar os benefícios dos *leaderboards* e amenizar seus possíveis malefícios.

### 3.1 Requisitos

**Requisito 1: Os *leaderboards* devem permitir mais oportunidades de vitória para mais alunos.**

É possível que, durante competições, exista um subconjunto de participantes que estão disputando as primeiras posições, enquanto os demais apresentam menos condições de alcançá-las. Esses participantes com um desempenho geral pior em relação aos seus pares podem ter chances praticamente nulas de vencer a competição. As experiências de vitória são positivas para o participante do contexto competitivo pois, de acordo com Reeve e Deci [43], vencer pode promover no participante senso de competência, aumentando assim sua

motivação.

Segundo Landers, Bauer e Callan [35], o *leaderboard* representa um caso em que há a promoção de uma gama de objetivos distintos que os seus participantes podem buscar (e.g. ser o primeiro colocado, estar entre os dez primeiros, estar melhor colocado que um determinado participante) e que não apresenta instruções de quais objetivos valem a pena ser alcançados. Além disso, os objetivos podem ser revisados no decorrer da competição [35]. Desse modo, um participante pode se engajar com objetivos mais difíceis se ganhar confiança de que pode alcançá-los ou pode se engajar com objetivos mais fáceis se perceber que não irá conseguir alcançar o objetivo definido anteriormente. A dificuldade do objetivo é importante pois, de acordo com Locke e Latham [39], objetivos mais difíceis incentivam mais esforço do que objetivos mais fáceis.

**Requisito 2: Os *leaderboards* devem recompensar esforço e não habilidade do aluno.**

De acordo com Landers e Landers [36], indivíduos que apresentam menos habilidade que os demais participantes de uma competição podem se tornar menos motivados no decorrer da competição devido à repetidas experiências de insucesso. Em contextos educacionais, os alunos apresentam diferentes ritmos de aprendizado e avanço [32]. Existem alunos que conseguem dominar os assuntos com maior facilidade e avançam em um ritmo mais rápido, enquanto existem alunos que apresentam mais dificuldades no domínio desses assuntos, necessitando de mais tempo e esforço para dominá-los. Em uma situação competitiva entre esses estudantes, naturalmente o aluno que avança com mais facilidade tem vantagem em relação aos alunos com mais dificuldade.

É comum nos jogos que objetivos mais difíceis rendam ao jogador uma quantidade de pontos maior que objetivos mais fáceis. Embora esse critério de pontuação possa incentivar o jogador a progredir em nível de dificuldade, em uma competição, esse critério de pontuação diferenciado reforça as diferenças entre os jogadores, de modo que os jogadores mais habilidosos terão muito mais chances de vencer do que jogadores menos habilidosos, que terão que se esforçar mais para obter a mesma quantidade de pontos que o aluno mais habilidoso.

**Requisito 3: Os *leaderboards* devem reduzir a visibilidade dos alunos.**

Os *leaderboards* exibem informações sobre o desempenho dos seus participantes de modo que cada participante pode ver o desempenho dos demais participantes e pode ter seu desempenho visualizado por estes participantes. A alta exposição de informação dos alunos pode representar um fator desmotivante para eles, como para aqueles que estão em posições ruins no *leaderboard* [8].

**Requisito 4: Os *leaderboards* devem realçar os alunos com melhor desempenho, como uma forma de lhes prestar reconhecimento.**

Por permitirem que o desempenho do participante seja visível por outras pessoas, os *leaderboards* podem ser utilizados como uma forma do participante demonstrar sua competência ou sua performance para os demais, podendo ser reconhecido pelos seus pares pelo seu esforço. Essa característica é interessante principalmente para os alunos nas posições do topo dos *leaderboards*, que correspondem aos alunos com performance melhor.

**Requisito 5: Os alunos devem participar da competição de forma voluntária.**

Uma das características dos jogos é que eles representam atividades voluntárias [4][41][48]. Ou seja, são atividades em que o indivíduo tem liberdade de participar ou não. A voluntariedade na participação é uma característica importante, visto que desse modo as influências externas são menores, havendo menos fatores controladores que iriam afetar negativamente o indivíduo [12], facilitando uma postura lúdica por parte dos jogadores.

É comum que existam disciplinas nas quais, para incentivar o aluno a resolver exercícios, se determina que a realização de exercícios tenha uma relação direta com a nota final do aluno. Essa abordagem faz com que se exista um fator controlador mais forte, pois o sucesso do aluno na disciplina passa a ter maior dependência da realização de exercícios. Essa estratégia possui impactos negativos como causar desconforto nos alunos, que podem se sentir obrigados a resolver os exercícios.

## 3.2 Decisões de design

**Decisão de design 1: *Leaderboards* de curta duração utilizados juntamente do *leaderbo-***

*ard geral.*

Essa decisão de design foi elaborada com o objetivo de cumprir o requisito de que deve existir mais oportunidades de vitória para mais alunos durante a disciplina. Os *leaderboards* de curta duração representam *leaderboards* que contabilizam a pontuação referente aos exercícios resolvidos pelos alunos em intervalos de tempo bem definidos dentro da disciplina, ignorando os exercícios resolvidos pelos alunos nos demais momentos. São exemplos de *leaderboards* desse tipo os *leaderboards* mensais, semanais e diários.

Em geral, disciplinas introdutórias de programação possuem meses de duração. Sendo assim, a implantação de *leaderboards* de curta duração aumenta significativamente a quantidade de competições existente dentro da disciplina e, conseqüentemente, as oportunidades de vitória dos alunos. Além disso, esses *leaderboards* são resetados sempre que os seus prazos se encerram, criando assim uma competição nova. Essa característica permite que mais alunos, além dos que estão em posições altas do *leaderboard* geral, possam vencer competições dentro da disciplina (por exemplo, um aluno que está em uma posição intermediária no *leaderboard* geral pode resolver muitos exercícios dentro de um período de tempo e, dessa forma, vencer nas competições menores).

**Decisão de design 2: Os exercícios de programação possuem a mesma pontuação, independente dos assuntos abordados por eles.**

Essa decisão de design foi elaborada com o objetivo de cumprir o requisito de que o critério de pontuação para os *leaderboards* devem recompensar esforço e não habilidade dos alunos, de modo que o desempenho do aluno na competição dependa do número de exercícios que ele resolve. Fazer com que os exercícios de programação possuam um valor constante reduz as disparidades existentes entre alunos com diferentes níveis de domínio do conteúdo e ritmos de avanço na disciplina. Com esse critério, um aluno que está mais atrasado em relação a um colega pode se esforçar na realização de exercícios dos assuntos que está estudando, tendo assim mais chances de competir com um aluno mais avançado.

É importante ressaltar que é necessário variedade de exercícios dos diferentes assuntos da disciplina para tornar esta decisão de design mais eficiente. Existindo essa variedade, os alunos terão exercícios para resolver por um intervalo de tempo maior, tempo este em que

irão se capacitando para avançar nos assuntos, liberando assim mais exercícios novos para fazer. Existindo uma variedade pequena de problemas, o aluno mais atrasado poderá passar pelo caso de conseguir fazer todos os exercícios disponíveis e acabar ficando sem exercícios novos, enquanto alunos mais avançados terão uma variedade maior de problemas.

**Decisão de design 3: Cada aluno visualiza nos *leaderboards* apenas os participantes nas posições do topo, os participantes empatados com ele e os participantes em posições vizinhas a dele.**

Essa decisão de design foi elaborada com o objetivo de cumprir dois requisitos levantados. O primeiro requisito é o de que os *leaderboards* devem realçar a visibilidade dos alunos nas posições do topo, representando o reconhecimento por seus esforços. O segundo diz respeito à restrição de visibilidade dos demais alunos no *leaderboard*, de modo que cada aluno que não esteja nas posições do topo seja visualizado apenas por alunos em posições vizinhas, isto é, em uma situação idêntica.

O critério de tornar visível para todos apenas os alunos nas posições do topo dos *leaderboards* serve de incentivo para os alunos que querem ser visíveis e reconhecidos pelos seus colegas, indicando para eles que, para alcançar esta visibilidade, eles precisam resolver exercícios. Por sua vez, para os demais alunos, a opção de ser visualizado apenas por alunos em situação idêntica à dele na competição serve como uma maneira de ocultar o seu desempenho, reduzindo a sua visibilidade.

**Decisão de design 4: Deve existir uma forma de registrar os melhores colocados nas competições menores.**

Essa decisão de design foi elaborada com o objetivo de cumprir com o requisito de que deve existir formas de reconhecimento do esforço dos alunos nas melhores posições dos *leaderboards*. Ela está relacionada principalmente com os *leaderboards* de curta duração, descritos na decisão de design 1.

Os *leaderboards* de curta duração possuem prazos definidos e são resetados quando estes prazos se encerram, dando início a novas competições. Não existindo um registro do desempenho dos melhores alunos ao final destas competições, os alunos não tomam conhe-

cimento de quais alunos tiveram um bom desempenho em quais competições. Isso reduz o reconhecimento desses alunos, podendo existir casos em que o próprio aluno não tem conhecimento de que venceu alguma das competições. Para evitar esse problema, essa decisão de design sugere que, antes dos *leaderboards* serem resetados, as informações referentes aos alunos nas melhores posições das competições sejam salvas e disponibilizadas aos alunos (por exemplo, através de um link).

**Decisão de design 5: As competições estão totalmente desvinculadas do processo avaliativo da disciplina.**

Essa decisão de design foi elaborada com o objetivo de cumprir o requisito de que a realização de exercícios por parte dos alunos deve ser voluntária. O intuito desta decisão é tornar a realização de exercícios e a participação nas competições existentes na disciplina práticas mais lúdicas. Isso significa que a nota do aluno não depende do desempenho dele nas competições existentes dentro da disciplina, indicando que, caso o aluno não queira participar das competições, a sua nota final na disciplina não será afetada.

### 3.3 Avaliação do design

Para avaliarmos as decisões de design para *leaderboards* propostas, elaboramos e executamos dois estudos. O primeiro deles diz respeito a um experimento controlado de duas horas de duração. O segundo estudo representa uma versão com mais decisões de design implementadas e com tempo de duração maior que o primeiro, sendo executado durante um período letivo de uma disciplina introdutória de programação. Nós coletamos e avaliamos dados referentes às interações dos alunos com um sistema de correção automático utilizado por eles em uma disciplina introdutória de programação.

## Capítulo 4

# Avaliação do efeito de um *leaderboard* na quantidade de atividade desempenhada por alunos em uma aula de programação

Nesta seção nós descrevemos e discutimos um estudo experimental. Nós investigamos se os *leaderboards* exercem efeito positivo nos alunos no que diz respeito à quantidade de exercícios realizados por eles durante uma aula prática. Após essa análise, investigamos os impactos da utilização de um critério de pontuação que pontua os exercícios igualmente no comportamento dos alunos.

Esse capítulo apresenta a seguinte estrutura:

- A seção 4.1 apresenta as perguntas de pesquisa que motivaram o experimento;
- A seção 4.2 descreve a estrutura e condução do experimento;
- A seção 4.3 traz os resultados encontrados;
- A seção 4.4 contém discussões levantadas;
- A seção 4.5 aborda as ameaças à validade.

## 4.1 Perguntas de pesquisa

O objetivo deste estudo foi o de avaliar de forma experimental a utilização de um *leaderboard* no que diz respeito às atividades realizadas pelos alunos durante uma aula prática de programação e implicações de decisões de design para estes *leaderboards*. Este estudo busca responder as seguintes perguntas de pesquisa:

**Pergunta de pesquisa 1:** A presença de *leaderboards* tem efeito na quantidade de atividade realizada pelos alunos em disciplinas introdutórias de programação?

Inicialmente, nós investigamos se a presença de um *leaderboard* desempenha um efeito positivo nos alunos. Para isso, nós decompomos o conceito de atividade e criamos algumas métricas para representá-la. Consideramos, para cada aluno, os registros de visualização de problemas e submissão de soluções para um sistema de correção automática, além do seu tempo de permanência em sala.

A partir desses registros, criamos as seguintes métricas:

**Quantidade de questões lidas (*READS*):** Representa a quantidade de enunciados de problemas diferentes visualizados pelo aluno durante o experimento. No contexto deste estudo, visualizar um enunciado é clicar em um link referente a um problema de programação. Nesse link, o aluno encontrava a descrição completa de um exercício, além de exemplos de entrada e saída para esse exercício.

**Quantidade de submissões realizadas (*SUBS*):** Representa a quantidade de submissões de código realizadas pelo aluno para o sistema de correção automática no experimento.

**Quantidade de problemas diferentes abordados (*PROBLEMS*):** Representa a quantidade de problemas diferentes que o aluno tentou resolver durante o experimento. Consideramos que um aluno tentou resolver um problema quando ele enviou para o sistema de correção automática, no mínimo, um código referente a esse problema, seja o código uma solução correta ou não segundo os testes de aceitação criados para o problema. Por conta

de limitações do sistema, que só coletava registros dos envios de código, não consideramos como sendo uma tentativa casos em que o aluno desenvolveu código para um problema sem enviá-lo para o sistema.

**Quantidade de problemas diferentes resolvidos (*SOLVED*):** Representa a quantidade de problemas diferentes que o aluno conseguiu resolver no experimento. Consideramos que o aluno resolveu um problema quando houve a submissão de um código que satisfaz todos os testes de aceitação do sistema de correção automático para esse problema. Essa métrica está diretamente relacionada com a quantidade de pontos que o aluno conseguiu durante o experimento. Cada problema resolvido corretamente por um aluno contava para ele 1 ponto.

**Tempo de permanência do aluno em sala de aula (*TIME*):** Representa o tempo de permanência voluntário do aluno em sala durante o experimento. Os alunos permaneceram em sala por no mínimo 90 minutos, devido a uma instrução dada pelos professores. Após esse tempo, os professores informaram aos alunos que eles poderiam permanecer por até 30 minutos a mais, se desejassem continuar exercitando. Por conta de limitações do sistema, que não permitia que soubéssemos exatamente o horário em que o aluno saiu do sistema, antes de sair da sala de aula, cada aluno preencheu uma lista de presença. Os professores ou assistentes presentes registraram manualmente o horário de saída dos alunos.

**Pergunta de pesquisa 2:** Quais os impactos da utilização de um critério de pontuação que pontua todos os exercícios igualmente?

Em geral, conforme o jogador avança em um jogo, os objetivos se tornam mais difíceis e rendem mais pontos. Esse critério de pontuação, entretanto, realça as diferenças de ritmo dos alunos, de modo que os alunos que apresentam um ritmo de avanço mais rápido criam uma vantagem mais acentuada em relação aos alunos com ritmo mais lento, que precisam exercer muito mais esforço para a obtenção da mesma quantidade de pontos.

Nesse trabalho, propomos um critério de pontuação que recompense o aluno pelo esforço e não pela sua habilidade, de modo que os alunos com ritmo de aprendizado mais lento tenham maiores chances de competir com alunos com ritmo mais rápido. Nós propomos um

critério de pontuação que considera que qualquer exercício, independente do assunto, rende aos alunos a mesma pontuação. Esse critério de pontuação foi utilizado no estudo e através dele investigamos como se dá o comportamento dos alunos e quais as chances que alunos mais atrasados tem na competição criada.

## 4.2 Metodologia

Nesta seção, iremos detalhar a metodologia seguida pelo experimento, informando detalhes sobre seus participantes, exercícios disponibilizados, design do *leaderboard* utilizado e sobre a condução do experimento.

### 4.2.1 Participantes

Nós tivemos a presença de 35 alunos novatos matriculados nas disciplinas de Programação I e Laboratório de Programação I do curso de Ciência da Computação da Universidade Federal de Campina Grande (UFCG). Destes 35 alunos, 18 participaram em um grupo de controle (i.e. grupo sem acesso ao *leaderboard*) e 17 em um grupo experimental (i.e. grupo com acesso ao *leaderboard*).

Cada aluno que participou do experimento tinha um nível de domínio diferente dos assuntos. Devido à filosofia de ensino *Mastery Learning* [6] utilizada nas disciplinas, que preza por um acompanhamento personalizado dos alunos, foi possível para os professores saberem o nível de domínio de cada estudante durante a disciplina. Esse domínio diferenciado foi levado em consideração nas análises dos resultados.

A separação dos alunos entre os grupos foi feita de forma aleatória. Para aumentar nossa confiança de que os grupos foram uniformemente divididos no que diz respeito ao domínio do conteúdo de seus participantes, fizemos um teste de hipótese de *Wilcoxon-Mann-Whitney* para verificarmos se os dois grupos eram estatisticamente equivalentes neste critério. O teste resultou no p-valor 0.37, que não rejeita a hipótese nula de que os grupos são equivalentes, reforçando nossa confiança de que não havia desequilíbrio entre os grupos no que se refere ao domínio do conteúdo dos alunos.

### 4.2.2 Questões

Nós selecionamos manualmente 55 questões pertencentes ao banco de questões do sistema de correção automática utilizado na disciplina. As questões utilizadas do experimento foram divididas da seguinte forma:

- 10 questões de entrada e saída e formatação de *strings* (unidade 2 da disciplina);
- 10 questões de blocos de código condicionais (unidade 3 da disciplina);
- 10 questões de iteração com o comando *for* (unidade 4 da disciplina);
- 10 questões de iteração com o comando *while* (unidade 5 da disciplina);
- 10 questões de definição e utilização de funções (unidade 6 da disciplina);
- 5 questões de estruturas de dados básicas (e.g. listas, filas) (unidade 7 da disciplina).

Essas questões foram exibidas no sistema de forma ordenada de acordo com a unidade que pertencem, seguindo uma ordem crescente de unidades. A ordem de exibição das questões foi a mesma para todos os alunos.

### 4.2.3 Leaderboard

Para o experimento, introduzimos um *leaderboard* com design simples, que foi projetado na parede da sala onde estava o grupo experimental. O *leaderboard* possuía as seguintes características:

- O *leaderboard* exibia três informações referentes ao aluno: a sua posição, o seu nome e a sua pontuação;
- A pontuação do aluno correspondia à quantidade de exercícios resolvidos por ele durante o experimento;

- A atualização do *leaderboard* foi realizada manualmente pelo professor ou assistente dentro da sala onde estava o grupo experimental, em média, uma vez a cada minuto. A decisão de atualizar manualmente foi tomada por conta de limitações do sistema, que não realizava uma atualização automática do *leaderboard*.

No experimento, o *leaderboard* exibiu a pontuação dos alunos de ambos os grupos. Essa decisão foi tomada por conta de limitações do sistema, pois não sabíamos de antemão quais alunos participariam do experimento nem em qual grupo participariam, e o sistema não realizava distinções em relação ao local de origem dos envios.

#### 4.2.4 Condução do experimento

O experimento foi realizado em uma segunda-feira e teve início às 10h da manhã. Na data do experimento, cerca de 70% do período letivo havia sido cumprido. Anteriormente à aula prática na qual foi realizada o experimento, os alunos estavam presentes na aula teórica da disciplina e não tinham, até então, conhecimento sobre a existência do experimento. Terminada a aula teórica, 35 alunos que estiveram na aula prática foram separados em duas turmas, sendo uma turma de controle com 18 alunos e uma turma experimental com 17 alunos. Cada turma possuía um professor e um estagiário docente da disciplina que atuaram como assistentes, ficando responsáveis pela condução do experimento. Eles também atuaram na fiscalização, impedindo que alunos entrassem em contato entre si e com alunos de outras turmas através de quaisquer fontes de comunicação, impedindo possíveis trocas de código ou informação sobre as questões entre os alunos.

Durante o experimento, os alunos interagiram com um ambiente controlado [23] que bloqueava acessos a *websites*, permitindo acesso apenas ao sistema de correção automática e o site oficial da documentação Python<sup>1</sup>. Desse modo, os alunos não puderam consultar sites externos para fazer pesquisa ou para se comunicar com outras pessoas, o que poderia afetar os resultados do experimento.

Inicialmente, os professores orientaram os alunos a realizarem *login* em suas contas no sistema. Em seguida, os professores informaram aos alunos que eles receberiam um con-

---

<sup>1</sup>Os alunos utilizaram a versão 2.7 da linguagem Python no experimento. A documentação desta versão está no link <https://docs.python.org/2.7/>

junto de questões do sistema e então os instruíram a tentar resolver o maior número possível de problemas dentro de 90 minutos como parte de um teste do sistema, que havia sido submetido a uma mudança fictícia. No grupo experimental, o professor menciona o *leaderboard* e explica como ele funciona, iniciando após a explicação a projeção na parede da sala. Após o início da projeção do *leaderboard*, os alunos de ambas as turmas começaram a programar. Para evitar enviesamentos, os professores e assistentes não tiraram dúvidas referentes aos exercícios ou à solução deles, limitando-se a responder perguntas referentes apenas à usabilidade do sistema e eventuais dúvidas acerca da condução do experimento.

Passados os primeiros 90 minutos do experimento, os professores informaram aos alunos que eles poderiam optar entre sair de sala ou continuar fazendo exercícios por até 30 minutos a mais. Quando o aluno decidia sair de sala, este preenchia uma lista de presença. Um professor ou assistente anotava o horário de saída desse aluno. Após os 30 minutos adicionais, os professores decretaram o fim da aula, encerrando assim a coleta de dados.

## 4.3 Resultados

### 4.3.1 Análise das métricas de atividade

Para a análise dos dados, utilizamos inicialmente estatística descritiva. Nós traçamos *boxplots* para cada métrica estudada, para cada grupo, além de calcularmos as médias e desvios padrão para estas métricas.

As Figuras 4.1, 4.2, 4.3, 4.4 e 4.5 exibem os *boxplots* de cada métrica do experimento.

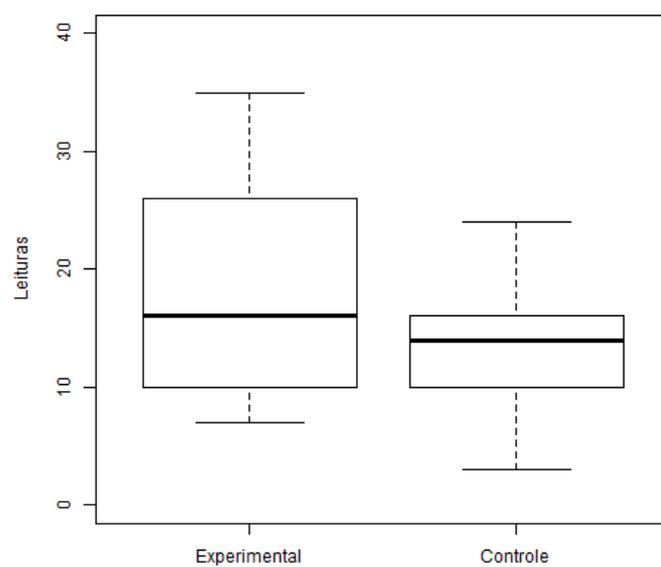


Figura 4.1: *Boxplot* da métrica *READS* entre os dois grupos

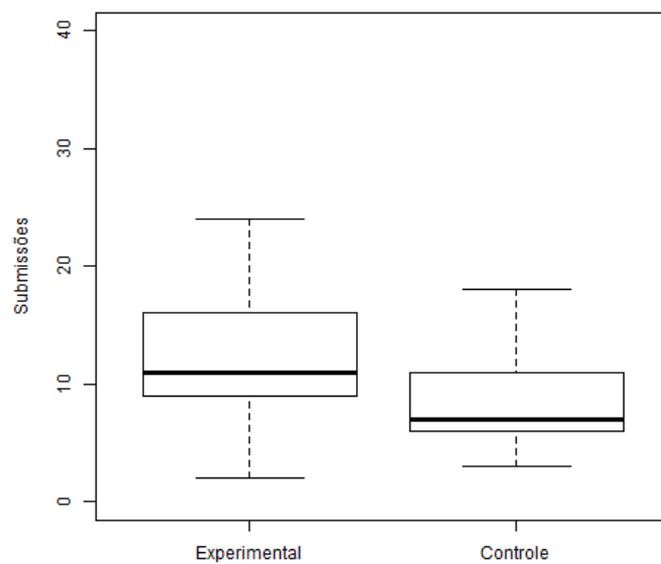


Figura 4.2: *Boxplot* da métrica *SUBS* entre os dois grupos

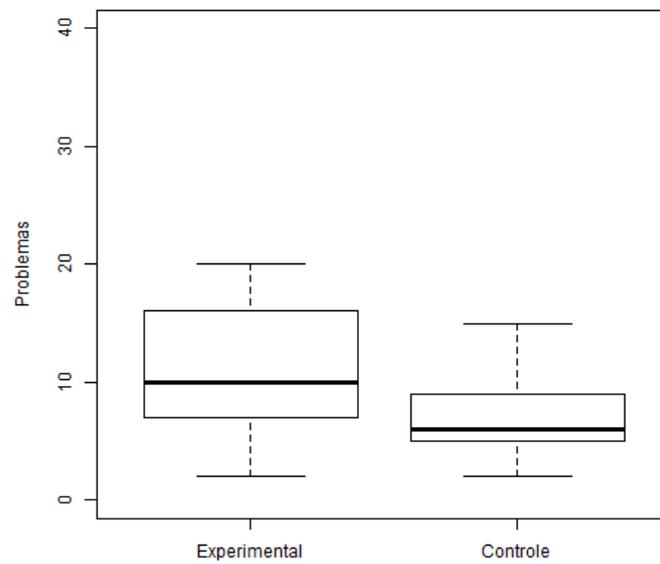


Figura 4.3: *Boxplot* da métrica *PROBLEMS* entre os dois grupos

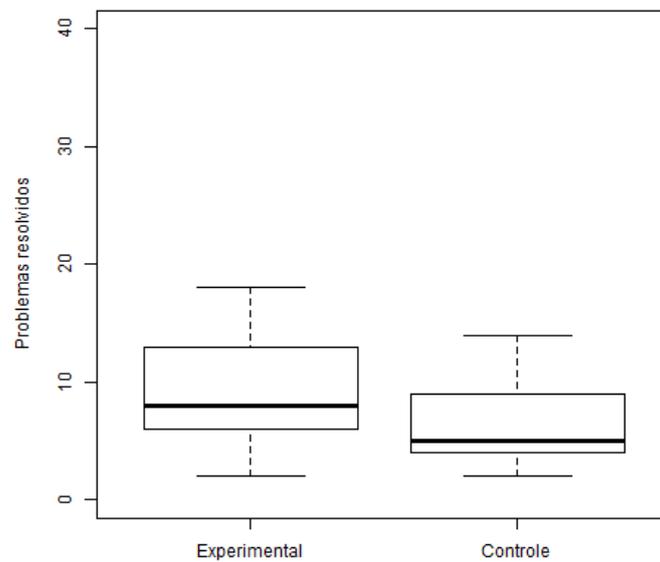


Figura 4.4: *Boxplot* da métrica *SOLVED* entre os dois grupos

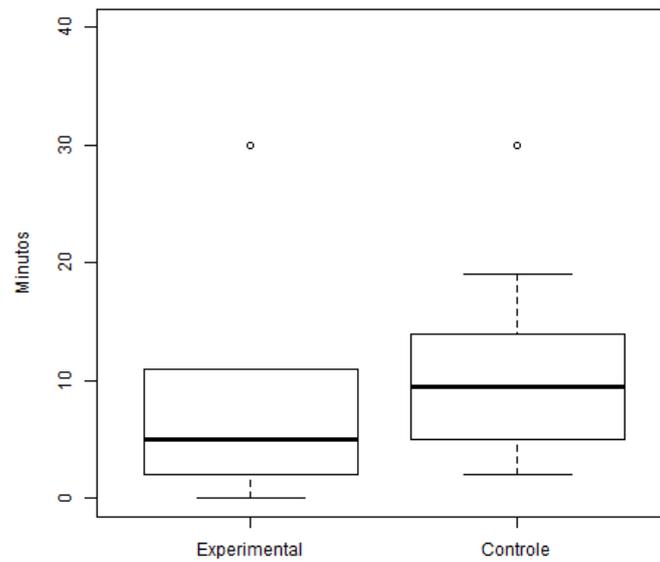


Figura 4.5: *Boxplot* da métrica *TIME* entre os dois grupos

As médias e desvios padrão das métricas estudadas para cada grupo podem ser encontradas nas Tabelas 4.1 e 4.2.

Tabela 4.1: Médias de cada métrica para cada grupo

Métrica	Média do grupo experimental	Média do grupo de controle
<i>READS</i>	18.059	12.889
<i>SUBS</i>	12.765	8.611
<i>PROBLEMS</i>	11.000	7.389
<i>SOLVED</i>	9.412	6.333
<i>TIME</i>	8.470	11.222

Tabela 4.2: Desvios padrão de cada métrica para cada grupo

Métrica	Desvio padrão do grupo experimental	Desvio padrão do grupo de controle
<i>READS</i>	8.525	5.635
<i>SUBS</i>	6.379	4.381
<i>PROBLEMS</i>	5.511	3.837
<i>SOLVED</i>	4.810	3.710
<i>TIME</i>	8.847	8.200

Os *boxplots* e os valores de desvio padrão indicam que a variabilidade dos dados é, em geral, maior no grupo experimental que no grupo de controle.

Para verificarmos diferenças estatísticas entre os valores das métricas entre os grupos, realizamos testes de hipótese de *Wilcoxon-Mann-Whitney*. A escolha do teste estatístico de *Wilcoxon-Mann-Whitney* se deu por conta das características dos nossos dados, que não seguem uma distribuição normal e são provenientes de grupos independentes. Os testes de hipótese foram feitos considerando as seguintes hipóteses:

H1.0: As medianas de *READS* são iguais entre os grupos.

H1.1: As medianas de *READS* são diferentes entre os grupos.

H2.0: As medianas de *SUBS* são iguais entre os grupos.

H2.1: As medianas de *SUBS* são diferentes entre os grupos.

H3.0: As medianas de *PROBLEMS* são iguais entre os grupos.

H3.1: As medianas de *PROBLEMS* são diferentes entre os grupos.

H4.0: As medianas de *SOLVED* são iguais entre os grupos.

H4.1: As medianas de *SOLVED* são diferentes entre os grupos.

H5.0: As medianas de *TIME* são iguais entre os grupos.

H5.1: As medianas de *TIME* são diferentes entre os grupos.

Para cada métrica, o teste resulta um p-valor que, se for menor que 0.05, rejeita a hipótese

nula, representando um indício de diferença de valores entre os grupos. A Tabela 4.3 contém os resultados dos testes de hipótese de *Wilcoxon-Mann-Whitney* para cada métrica criada para este estudo.

Tabela 4.3: p-valores dos testes de hipótese para cada métrica entre os grupos

Métrica	p-valor
<i>READS</i>	0.120
<i>SUBS</i>	0.038
<i>PROBLEMS</i>	0.028
<i>SOLVED</i>	0.042
<i>TIME</i>	0.123

Para *SUBS*, *PROBLEMS* e *SOLVED* os p-valores obtidos foram menores que 0.05, nos dando indícios de que os valores para estas métricas são diferentes entre o grupo de controle e o experimental. Para essas três métricas, o grupo experimental obteve valores maiores que o grupo de controle, o que indica que o *leaderboard* teve efeito positivo sobre elas. Por outro lado, para as métricas *READS* e *TIME*, a hipótese nula não foi rejeitada. Deste modo, não encontramos evidência de que o *leaderboard* teve efeito nessas métricas baseando-se nos testes de hipótese.

### 4.3.2 Tentativas de problemas por unidade

Fizemos também análises de acordo com as unidades dos problemas exercitados pelos alunos, com o objetivo de verificarmos o impacto do critério de pontuação quanto aos tipos de exercícios que os alunos se dedicaram no experimento. Nessa análise, não nos preocupamos se os alunos resolveram ou não os exercícios e sim se eles tentaram resolver, isto é, dedicaram tempo e esforço na resolução deles. Por conta disso, consideramos apenas os valores para a métrica *PROBLEMS* de cada estudante, que foram então separados de acordo com a unidade do problema abordado. Realizamos, então, um somatório destes valores para cada grupo. O resultado deste somatório pode ser encontrado na Tabela 4.4.

Houve uma maioria numérica em ambos os grupos quanto a codificação de exercícios das unidades 2 e 3, as unidades mais básicas dentre as consideradas no experimento. Esse

Tabela 4.4: Somatório de problemas abordados por unidade por alunos

Unidade	Grupo experimental	Grupo de controle
2	81	49
3	56	32
4	15	9
5	20	18
6	15	25
7	0	0

resultado era esperado de antemão, visto que mais alunos estão aptos a resolver exercícios dessas unidades do que das demais.

Após observarmos as quantidades de problemas abordados por unidade, procuramos também verificar em quais dessas unidades houveram diferença estatística entre os grupos. Para verificarmos essa diferença, realizamos testes de *Wilcoxon-Mann-Whitney* entre os valores por unidade de cada grupo, cujos resultados podem ser encontrados na Tabela 4.5. Devido ao fato de nenhum aluno ter ao menos tentado um exercício da unidade 7, essa unidade foi desconsiderada da análise.

Tabela 4.5: p-valores dos testes de hipótese de problemas abordados por unidade

Unidade	p-valor
2	0.020
3	0.028
4	0.516
5	0.627
6	0.313

Podemos dizer que, embora tenhamos encontrado evidências de que o *leaderboard* incentivou os alunos a abordarem mais problemas, esse incentivo ocorreu principalmente nas unidades 2 e 3, as mais básicas dentre as consideradas no experimento. Para as unidades 4, 5, 6 e 7, não encontramos evidência de que há diferenças estatísticas entre os grupos. Esse resultado sugere que o *leaderboard* e o critério de pontuação que atribui o mesmo va-

lor para todos os exercícios desempenharam um efeito aditivo, incentivando a realização de exercícios das unidades iniciais sem reduzir a quantidade de atividade abordada nas demais unidades.

Nós investigamos os comportamentos dos alunos individualmente quanto aos problemas tentados por eles no experimento. As Tabelas 4.6 e 4.7 exibem as unidades dos alunos, respectivamente, do grupo experimental e do de controle, e as unidades dos problemas tentados por eles na ordem em que foram tentados.

Tabela 4.6: Unidades dos problemas tentados por alunos do grupo experimental

Unidade do aluno	Unidade dos problemas tentados por ele
3	4,2
4	2,3,2,3,2,2
4	2,3,2,3,3,3
4	2,4,3,2,2,5,3,2,2,2,2,3,3,3,3
5	2,2,2,2,3,2,2,3,3
5	2,5,2,2,2,3,3,6,2
5	4,3,2,2,5,2,3,5,4,2,2,6,3,4,3,3,2
5	6,2,4,3,5,2,3,2,4,2,6,6,4
6	2,2,2,3,2,2,3
6	2,2,2,3,2,2,2,3,3,2
6	2,3,2,6,2,2,2,2,2,5,3,3,4,3,2
6	3,2,5,3,5,2,2,2,2,3,3,3,3
6	2,2,5,5,2,3,2,3
6	6,2,3,2,2,5,2,3,2,4,2,3,6,3,3,6
6	6,4,4,2,2,5,5,3,5,4,5,3,6,3,3,3,2,2,6,4
6	5,5
7	2,2,2,2,5,5,5,3,6,4,2,3,6,3,3,3,2,6

Tabela 4.7: Unidades dos problemas tentados por alunos do grupo de controle

Unidade do aluno	Unidade dos problemas tentados por ele
3	2,2,2,3,2,3,2
3	3,2,2,3,2,2,3
4	2,5,5,4
5	6,2,3,2,3,2
5	4,2,3,6,4
5	6,5,5,5,5
6	2,2,2,3,2,2,3,3,3
6	4,2,2,6,5,3,2,2,6,3,3,3,2
6	2,3
6	2,2,4,2,3,5,4,2,2,2
6	2,2,6,2,3,5,4,2,2,6,3,3,3,3,3
6	5,6,4,2
6	5,6,5,5,2,6
6	6,4,5,2,6,6
6	6,2,6,6,5
7	2,2,3,2,3,5,2,2,2,2,3,3,3,3,2
7	6,6,6,3,2
7	6,6,5,5,6,6,6,6

É possível percebermos que existiram nos dois grupos alunos que se concentraram na abordagem de exercícios de assuntos básicos, mesmo que esses alunos tivessem domínio de muitos conteúdos da disciplina. Além disso, os alunos pareceram não seguir uma ordem de resolução de problemas, constantemente retomando exercícios de unidades anteriores. Por fim, é possível observar que poucos alunos tentaram exercícios de unidades além da unidade em que eles estavam no dia do experimento.

### 4.3.3 Posição dos alunos

Na Tabela 4.8, disponibilizamos informações sobre os alunos do topo do *leaderboard* no final da aula. Nós enfatizamos os alunos do topo com o intuito de observarmos se o *leaderboard* utilizado permitiu alunos mais atrasados em relação a seus colegas conseguirem alcançar altas posições.

Tabela 4.8: Resultado final do *leaderboard*

Posição	pontos	grupo	domínio
1º	18	E	7
2º	16	E	6
2º	16	E	6
4º	14	E	4
4º	14	C	6
6º	13	E	5
6º	13	C	7
8º	12	E	6
8º	12	C	6
10º	10	E	6
10º	10	E	5
12º	9	C	6
12º	9	C	6
14º	8	E	6
14º	8	E	5
14º	8	C	7
17º	7	E	6
17º	7	E	5

Podemos perceber que existem alunos do grupo experimental que estão atrasados em relação a seus colegas que alcançaram altas posições do *leaderboard* (um na 4ª posição e um na 6ª posição). Esses alunos compensaram o domínio reduzido de assuntos realizando mais exercícios sobre os assuntos que eles dominavam.

## 4.4 Discussões

Nós iremos nesta seção discutir com mais detalhes os resultados encontrados no experimento. Nós planejamos, elaboramos e implementamos um design de *leaderboards* a ser utilizado em disciplinas de programação introdutória. Nesse experimento, investigamos se os *leaderboards* tem efeito na quantidade de atividade desempenhada pelos alunos, além de verificarmos os impactos da adoção de um critério de pontuação que considera que todos os exercícios possuem a mesma pontuação, independente de sua dificuldade.

### 4.4.1 Análise das métricas de atividade

#### Métrica *READS*

Nós investigamos se a existência de um *leaderboard* faria com que os alunos visualizassem mais enunciados de problemas diferentes. O objetivo dessa investigação foi verificar se o *leaderboard* incentivaria o aluno a visualizar mais enunciados, com o objetivo de procurar por exercícios mais fáceis na concepção dele e então concentrar seus esforços nesses exercícios. A análise dos dados obtidos, entretanto, não nos permite observar indícios de que este fenômeno tenha ocorrido.

#### Métrica *SUBS*, *PROBLEMS* e *SOLVED*

Foi possível observar que os alunos do grupo com o *leaderboard* submeteram mais códigos para mais problemas diferentes do que o grupo de controle, o que indica que a presença do *leaderboard* pode tornar os alunos mais ativos. Este aumento de atividade foi acompanhado de um aumento no número de problemas resolvidos, mostrando que os alunos apresentaram um melhor desempenho e rendimento com os *leaderboards*.

É importante ressaltar que existe efeito positivo do *leaderboard* nas três métricas referentes a implementação de código. Um crescimento apenas em *SUBS* indicaria que a presença do *leaderboard* faz com que os alunos realizem mais submissões incorretas para o sistema, o que significaria que o *leaderboard* pode incentivar o aluno a realizar submissões de forma menos consciente ou apressada. Um crescimento em *SUBS* e *PROBLEMS* iria sugerir que os alunos tentam mais problemas, mas aumentam o número de questões em aberto (i.e. sem solução), o que não é desejável.

### **Métrica TIME**

Quanto ao tempo de permanência dos alunos, não encontramos evidências de diferença entre os dois grupos. Nós acreditamos que 90 minutos programando corresponda a um período de tempo longo para os alunos, de modo que parte deles tenha se sentido desencorajada a continuar programando mesmo com a existência de um *leaderboard*. A existência de *leaderboards* pode ter feito alguns alunos cansarem mais rápido, o que justificaria os valores menores de tempo de permanência dos alunos, embora essa diferença não tenha sido encontrada estatisticamente. Entretanto, esse resultado pode ser interpretado de forma positiva, visto que, existindo uma possível igualdade em tempo de permanência estatístico entre os dois grupos, temos uma evidência de que a presença do *leaderboard* faz com que os alunos exercitem mais dentro de um mesmo período de tempo.

#### **4.4.2 Tentativas de problemas por unidade**

A análise dos resultados indica que o *leaderboard* com um critério de pontuação que determina que todos os exercícios possuem o mesmo valor incentiva a realização de exercícios envolvendo os assuntos mais simples dentre os considerados, sem que tenhamos observado uma redução na realização de exercícios de assuntos mais avançados.

Pelo fato de que os assuntos são cumulativos em disciplinas introdutórias de programação, os exercícios de assuntos mais simples possuem descrições menores, cobram menos conteúdo e exigem menos código dos alunos. Por conta dessas características, exercícios dessas unidades tendem a ser respondidos com maior rapidez, rendendo ao aluno mais pontos em menos tempo dentro da competição. Nós observamos que existem casos de alunos que realizam exercícios apenas das unidades básicas, mesmo apresentando domínio de muitas unidades da disciplina. Esse comportamento ocorreu em alunos dos dois grupos e é necessário uma maior investigação para verificar se os *leaderboards* com pontuação igual entre questões estimulam o aluno a focar em assuntos anteriores ao que ele deveria estar estudando.

Vale ressaltar que esse experimento foi conduzido próximo ao final do período letivo. Nesse caso, a estratégia de fazer exercícios das unidades anteriores é muito benéfica para o aluno na competição, visto que é uma forma de conseguir pontos com menos esforço. Nós

acreditamos que esse efeito é mitigado se os *leaderboards* forem disponibilizados aos alunos no momento em que eles começam a resolver exercícios dentro da disciplina. No começo da disciplina, o aluno tem acesso a um conjunto pequeno de assuntos e irá fazer exercícios sobre esses assuntos. Conforme avança nos assuntos da disciplina, o aluno irá avançar também nos exercícios. Assim, há uma aproximação maior entre o conteúdo que está sendo estudado pelo aluno e o conteúdo dos exercícios que o mesmo resolve.

### 4.4.3 Posição dos alunos

No experimento, disponibilizamos aos alunos 55 exercícios sobre diferentes assuntos da disciplina. Além disso, os exercícios possuíam o mesmo valor, independente dos assuntos que o exercício abordava. Esses dois pontos tiveram como objetivo permitir que, na competição, os alunos com menor domínio de conteúdos da disciplina pudessem ter mais chances de competir com alunos de domínio maior.

Nós observamos que existiram alunos mais atrasados em relação a seus colegas que conseguiram alcançar altas posições dentro do *leaderboard*. Esses alunos compensaram a menor variedade de assuntos que poderiam abordar no experimento com a realização de exercícios sobre os assuntos que eles dominavam. Este resultado é um indicativo positivo de que pontuar os exercícios com o mesmo valor permite com que alunos com diferentes níveis de domínio alcancem boas posições no *leaderboard*, o que pode motivar alunos com domínio menor a realizarem exercícios.

## 4.5 Ameaças à validade

Uma ameaça à validade do experimento está relacionada com o tamanho da amostra utilizada. A quantidade de participantes do experimento foi de 18 alunos em um grupo de controle e 17 em um grupo experimental. Essa quantidade de participantes é pequena, o que não nos permite obter resultados mais poderosos no ponto de vista estatístico. São necessárias mais replicações desse experimento para que os resultados sejam mais significativos.

Os professores de ambas as turmas pediram que os alunos resolvessem o maior número de exercícios que conseguissem dentro de uma aula. Essa instrução pode ter tido efeito nos alunos, que se sentiram incentivados a realizar exercícios devido à ela. No grupo em que

---

os alunos tinham acesso aos *leaderboards*, pode existir confusão entre esses dois fatores. Deste modo, é possível que no grupo experimental, existam alunos que não foram totalmente incentivados pelos *leaderboards*.

# Capítulo 5

## Avaliação da utilização de *leaderboards* durante um período letivo

Neste capítulo, apresentamos um estudo longitudinal com o objetivo de investigar os impactos da utilização de um design de *leaderboards* e a relação deles com a resolução de exercícios por parte dos alunos. Nós coletamos dados referentes à visualização de *leaderboards* e resolução de exercícios por parte dos alunos durante 70 dias de um período letivo de uma disciplina de programação introdutória. Esse capítulo apresenta a seguinte estrutura:

- A seção 5.1 apresenta as perguntas de pesquisa deste estudo;
- A seção 5.2 descreve a estrutura do estudo;
- A seção 5.3 traz os resultados encontrados;
- A seção 5.4 contém discussões levantadas.

### 5.1 Perguntas de pesquisa

O objetivo deste estudo foi avaliar decisões de design para *leaderboards* no que diz respeito à atividades resolvidas pelos alunos no decorrer de um período letivo. Esse estudo busca responder as seguintes perguntas de pesquisa:

**Pergunta de pesquisa 3:** A implantação de *leaderboards* de curta duração incentiva os alunos a exercitarem?

Em geral, competições apresentam disparidades entre a performance de seus participantes, de modo que um número considerável deles não possui chances de vencer ou mesmo alcançar altas posições nos *leaderboards*. Com o objetivo de reduzir esse problema, implantamos *leaderboards* de curta duração. No design de *leaderboards* proposto utilizamos, em conjunto com o *leaderboard* geral, que considera as atividades realizadas durante toda a disciplina, *leaderboards* diários e semanais. Esses *leaderboards* consideram as atividades realizadas pelo aluno, respectivamente, no dia e semana correntes.

A utilização de *leaderboards* de curta duração aumentam a quantidade de competições existentes, de modo a aumentar as chances de vitória para os alunos. Por contabilizarem os exercícios resolvidos pelos alunos durante intervalos de tempo curtos, desconsiderando os exercícios resolvidos anteriormente, os *leaderboards* semanais e diários favorecem alunos que estão em desvantagem em relação a seus colegas no cenário geral. Além disso, o critério de pontuação considerado pelos *leaderboards*, que pontua igualmente os exercícios independente dos assuntos abordados por eles, colaboram com uma competição mais equilibrada. Essas duas características permitem mais alunos terem chances de vitória.

**Pergunta de pesquisa 4:** Qual o nível de interesse dos alunos em relação aos *leaderboards* no decorrer de um período letivo?

A revisão da literatura de Hamari, Koivisto e Sarsa [25] aponta que o efeito da gamificação tende a ser reduzido com o tempo. Coletamos os registros de visualização dos *leaderboards* feitos pelos alunos no decorrer do período letivo com o objetivo de observar se esta tendência ocorre em uma disciplina introdutória de programação.

**Pergunta de pesquisa 5:** Qual a relação entre visualização dos *leaderboards* e resolução de exercícios por parte dos alunos?

Nós investigamos as visualizações dos *leaderboards* e a realização de exercícios feitos considerando toda a turma e casos particulares, com o objetivo de verificar possíveis traços de interesse manifestados pelos alunos em relação aos *leaderboards*.

## 5.2 Metodologia

Nesta seção, iremos detalhar a metodologia seguida pelo estudo, informando detalhes sobre seus participantes, métricas consideradas e o design de *leaderboards* utilizado.

### 5.2.1 Participantes

Foram analisados dados relativos ao segundo semestre letivo do ano de 2015 das disciplinas introdutórias de Programação do curso de Ciência da Computação da Universidade Federal de Campina Grande (UFCG). No semestre em questão, tivemos 112 alunos matriculados, dentre os quais 96 alunos foram considerados. Como critério de exclusão para o estudo, excluímos os alunos que, após a implantação dos *leaderboards*, não visualizaram nenhum dos *leaderboards* e não resolveram nenhum exercício do sistema de correção automático da disciplina.

Utilizamos os registros de atividade dos alunos gerados através da interação deles com um sistema de correção automática durante a disciplina entre os dias 11 de Fevereiro até o dia 29 de Maio de 2016, totalizando 109 dias. Os alunos utilizaram esse sistema para terem acesso aos exercícios de programação, além de poderem enviar códigos referentes aos exercícios para o sistema, que então fornece um *feedback* sobre esses códigos para o aluno.

### 5.2.2 Métricas

Consideramos, para cada aluno, duas métricas: a quantidade de visualizações dos *leaderboard* feitas e a quantidade de exercícios resolvidos. Chamamos estas métricas, respectivamente, de *VIEWS* e *SOLVED*.

A métrica *VIEWS* corresponde à quantidade de vezes que o aluno visualizou os *leaderboards* durante a disciplina. Essa métrica é baseada nos registros de visualização dos *leaderboards* por parte do aluno no sistema. Nesse trabalho, utilizamos três *leaderboards* diferentes:

- *day*: *leaderboard* que conta apenas os pontos obtidos no dia;
- *week*: *leaderboard* que conta apenas os pontos obtidos na semana;

- *full: leaderboard* que conta os pontos obtidos desde o começo da disciplina.

A métrica *SOLVED* corresponde à quantidade de exercícios disponíveis no sistema de correção automática que foram resolvidos pelo aluno durante a disciplina. Essa métrica é baseada nos registros de exercícios resolvidos pelo aluno no sistema. Um novo registro é gerado toda vez que o aluno resolve um novo problema.

### 5.2.3 Design dos *leaderboards*

Para este estudo, ampliamos o design de *leaderboards* utilizado em relação ao design do capítulo 4. O novo design possuiu as seguintes características.

- Foram utilizados, além de um *leaderboard* geral, mais dois tipos de *leaderboard*: diário e semanal, que consideram, respectivamente, os pontos obtidos pelos alunos no dia atual e na semana atual. Consideramos que um dia se inicia às 00:00 e termina às 23:59 e uma semana inicia na sexta-feira de 10:00 até sexta-feira de 08:00. Esses horários de início e fim da semana foram baseados nas avaliações semanais realizadas na disciplina, que são conduzidas nas sextas-feiras, de 08:00 até 10:00;
- Os *leaderboards* exibiam três informações referentes à cada aluno: a sua posição, o seu nome e a sua pontuação;
- Os *leaderboards* não exibiam para os alunos informações sobre todos os estudantes. Os alunos que estavam nas primeiras cinco posições dos *leaderboards* eram exibidos para todos. Para cada aluno, o *leaderboard* exibia os alunos que estavam empatados com ele, até duas posições acima dele e até duas posições abaixo dele. As informações sobre os demais alunos não eram visíveis;
- O cálculo da pontuação do aluno nestes *leaderboards* foi baseado apenas na quantidade de exercícios resolvidos pelo aluno durante a disciplina.
- A atualização do *leaderboard* era realizada sempre que a pontuação de um aluno mudava (i.e. um aluno resolvia um exercício não resolvido anteriormente) pelo sistema de correção automática.

- A realização de exercícios esteve totalmente desvinculada do processo avaliativo da disciplina.

Os *leaderboards* foram inicializados no dia 21 de Março, 40 dias após o primeiro registro de submissão de código para o sistema de correção automático. Decidimos ligá-los após um certo tempo pois, nos primeiros dias, os alunos ainda estavam se adaptando ao sistema, além de que os *leaderboards* estavam tendo sua implementação finalizada. Os *leaderboards* permaneceram ligados por 70 dias do período letivo. Para os dias 15, 16 e 17 de abril (respectivamente, uma sexta-feira, um sábado e um domingo), não obtivemos todos os dados de visualização dos *leaderboards* devido à problemas na coleta dos registros. As visualizações nestes três dias foram completamente desconsideradas do estudo.

## 5.3 Resultados

Nós realizamos três análises: na primeira, investigamos a quantidade de visualizações dos *leaderboards* separadas por tipo, além de possíveis preferências dos alunos quanto a um tipo específico; na segunda, procuramos observar a quantidade de visualizações dos *leaderboards* e a resolução de exercícios divididas por dia, verificando estas quantidades no decorrer do tempo; por último, relacionamos as visualizações dos *leaderboards* e a resolução de exercícios para cada aluno.

### 5.3.1 Visualização dos *leaderboards* por tipo

A primeira análise consiste nas visualizações feitas pelos alunos divididas por tipo de *leaderboard*. As quantidades de visualizações dos *leaderboard* feitas, separadas por tipo, podem ser encontradas na Tabela 5.1.

Tabela 5.1: Quantidade de visualizações dos *leaderboards* separadas por tipo

Tipo do <i>leaderboard</i>	Visualizações obtidas	Porcentagem	Média	Desvio padrão
geral ( <i>full</i> )	3085	44.8%	32.47	44.27
semanal ( <i>week</i> )	1884	27.3%	19.81	25.77
diário ( <i>day</i> )	1921	27.9%	20.22	26.81

As Figuras 5.1, 5.2 e 5.3 representam, respectivamente, as visualizações ao *leaderboard* geral, semanal e diário de cada aluno. Os alunos são representados pelos pontos do gráfico e estão ordenados de forma decrescente de acordo com a quantidade de exercícios resolvidos durante a disciplina.

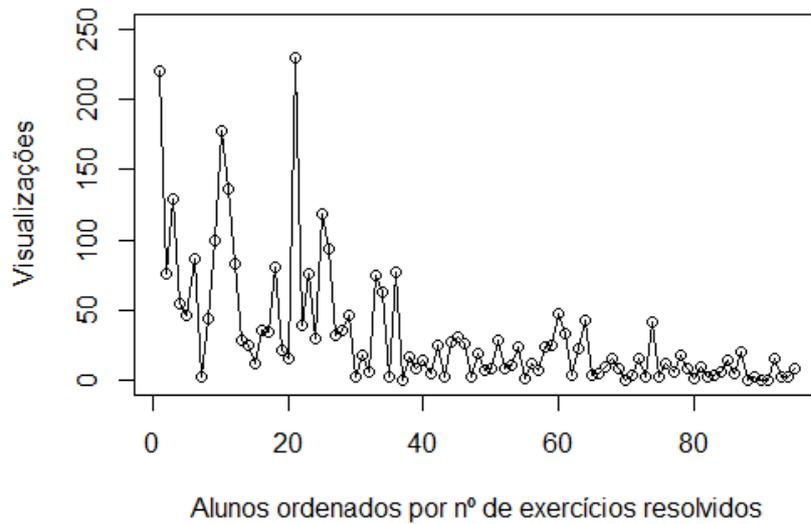


Figura 5.1: Quantidade de visualizações do *leaderboard* geral feitas pelos alunos

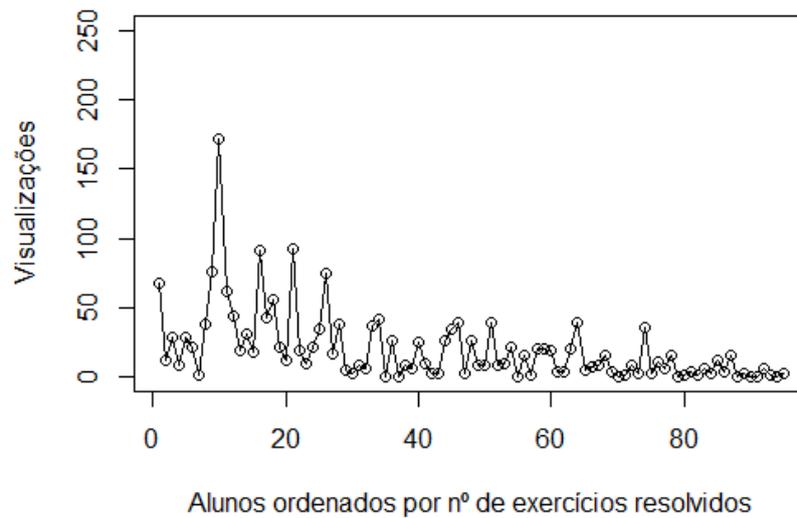


Figura 5.2: Quantidade de visualizações do *leaderboard* semanal feitas pelos alunos



Figura 5.3: Quantidade de visualizações do *leaderboard* diário feitas pelos alunos

Podemos perceber que os *leaderboards* são visualizados, principalmente, pelos alunos mais ativos da disciplina. Sobre esses alunos, é possível observar que eles visualizam o *leaderboard* geral mais do que os *leaderboards* menores. Em contrapartida, alunos que fazem poucos exercícios costumam apresentar uma quantidade de visualizações pequena, mesmo dos *leaderboards* menores.

Nós vimos que alguns alunos que resolveram uma quantidade de exercícios intermediária em relação à seus colegas apresentaram quantidade de visualizações dos *leaderboards* diário ou semanal maior do que do *leaderboard* geral. Através de uma investigação mais

aprofundada do histórico de atividades desses alunos, percebemos que essa diferença ocorre devido ao fato de que, em alguns dias, esses alunos acessaram os *leaderboards* menores e fizeram exercícios com maior frequência, de forma alternada. Esse aumento e alternância de ações representa um indício de aumento de engajamento, mesmo que por pouco tempo.

### 5.3.2 Visualização dos *leaderboards* e resolução de exercícios no tempo

A segunda análise diz respeito à quantidade de visualizações dos *leaderboards* e resoluções de exercícios feitas pelos alunos no decorrer do tempo. Para isto, contabilizamos o total de visualizações feitas e o total de exercícios resolvidos em cada dia. A quantidade de visualizações dos *leaderboards* feitas por dia, somando-se todos os tipos, pode ser consultada na Figura 5.4.

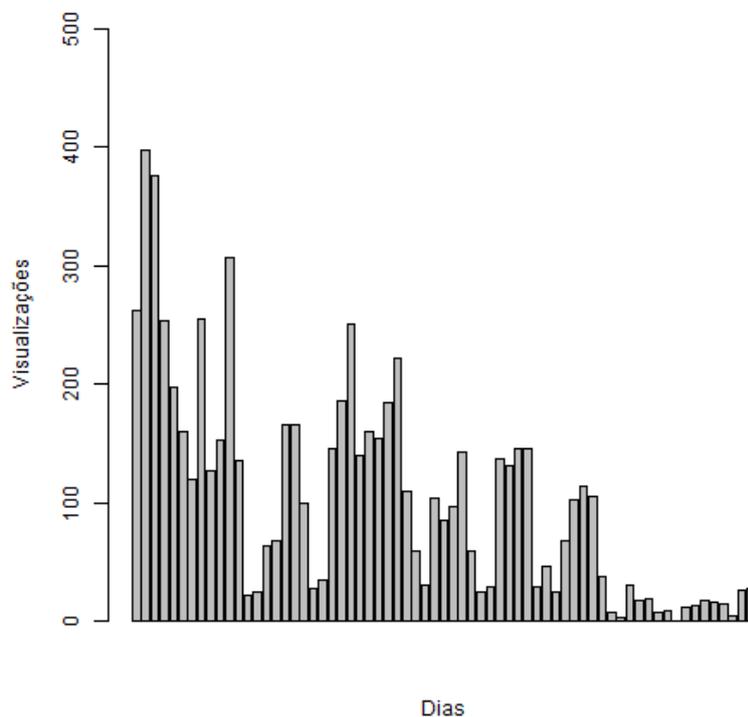


Figura 5.4: Quantidade de visualizações dos *leaderboards* feitas por dia pelos alunos

A Figura 5.5 traz a quantidade de exercícios resolvidos pelos alunos durante a disciplina, incluindo os dias em que os *leaderboards* ainda não tinham sido implantados no sistema de correção automática. A linha vertical preta da figura representa o dia em que os *leaderboards* foram disponibilizados para os alunos.

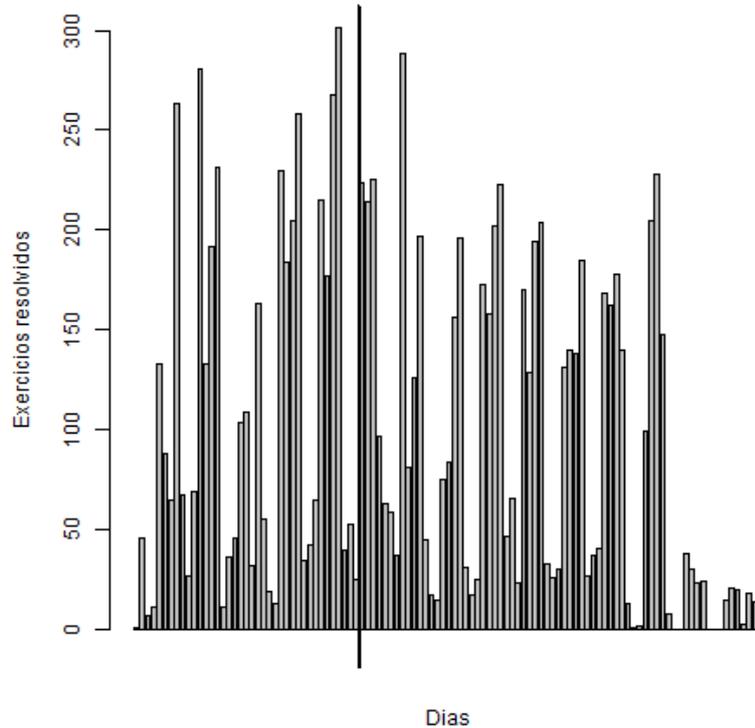


Figura 5.5: Quantidade de exercícios resolvidos pelos alunos em todos os dias

A alta variabilidade nestes dados é proveniente das constantes oscilações existentes na quantidade de atividade desempenhada pelos alunos durante as semanas. As quedas no número de visualizações de *leaderboards* e de resoluções de exercícios, em geral, ocorrem nos fins de semana, momento em que os alunos interagem menos com o sistema. Os aumentos ocorrem, em geral, nas segundas-feiras, quando os alunos retomam as suas atividades.

Nos primeiros dias, os *leaderboards* receberam muitos acessos e, com o passar do tempo, passaram a ser menos visualizados pelos alunos. A queda nos números com o passar do tempo quanto ao número de exercícios resolvidos diariamente pelos alunos, por sua vez, é menos acentuada.

Podemos observar que, nos últimos dias, a quantidade de visualizações dos *leaderboards* e de resoluções de exercícios realizados pelos alunos é bastante baixa em relação a dias an-

teriores. Esses dias representam o período de véspera das últimas avaliações semestrais dos alunos. Nesses dias, os alunos são submetidos a trabalhos e avaliações de outras disciplinas do período letivo e, por conta disso, passam a ter menos tempo e foco na realização de exercícios de programação.

### 5.3.3 Relações entre visualização de *leaderboards* e resolução de exercícios

Nesta seção, iremos verificar a relação entre as visualizações dos *leaderboards* e a quantidade de exercícios feitos pelos alunos. Nós calculamos a correlação de Spearman entre essas ações para cada aluno, além de investigarmos possíveis padrões de visualização e resolução por parte dos estudantes ao longo da disciplina.

Para calcularmos a correlação entre visualização de *leaderboards* e resolução de exercícios nós consideramos, para cada estudante, o intervalo de tempo em que ele interagiu com o sistema de correção automática, resolvendo exercícios ou visualizando os *leaderboards*. Esse intervalo se inicia no dia 21 de Março e se encerra no último dia em que o aluno fez alguma resolução de exercício ou visualização de *leaderboards*. Fizemos esse tratamento pois existem, na disciplina introdutória na qual o estudo foi conduzido, alunos desistentes e alunos que cumprem com o conteúdo da disciplina antes do tempo e, em geral, esses alunos não interagem com o sistema nessas condições.

Depois nós verificamos, para cada aluno, os dias em que ele visualizou um *leaderboard* e os dias em que ele resolveu um exercício. Feito isto nós calculamos, para cada aluno, a correlação de Spearman entre os dias em que o aluno visualizou algum *leaderboard* e os dias em que resolveu algum exercício, obtendo assim o valor de  $\rho$  para cada aluno.

A Tabela 5.2 organiza os alunos de acordo com os seus valores de  $\rho$ , obtidos com o cálculo da correlação de Spearman. A Tabela 5.3 exhibe, para cada aluno, a quantidade de dias em que ele visualizou um *leaderboard*, a quantidade de dias em que ele resolveu um exercício após a implantação dos *leaderboards*, o número de exercícios que ele resolveu antes dos *leaderboards*, o número de exercícios que resolveu após a implantação dos *leaderboards* e o seu valor de  $\rho$ .

Tabela 5.2: Correlação entre visualização de *leaderboard* e resolução de exercício dos alunos

Faixa do valor de $\rho$	Total de alunos	Porcentagem
menor que 0	4	4.2%
entre 0.0 e 0.2	4	4.2%
entre 0.2 e 0.4	23	24.0%
entre 0.4 e 0.6	36	37.5%
entre 0.6 e 0.8	21	22.0%
entre 0.8 e 1.0	2	2.1%

Tabela 5.3: Dados sobre visualização de *leaderboards* e resolução de exercícios dos alunos

Aluno	Dias que viu <i>leaderboard</i>	Dias que resolveu exercícios	Exercícios resolvidos antes dos <i>leaderboards</i>	Exercícios resolvidos após os <i>leaderboards</i>	Correlação entre visualização e resolução
1	7	9	27	28	0.860
2	13	10	110	44	0.849
3	20	21	26	66	0.791
4	22	27	56	108	0.774
5	24	21	36	133	0.764
6	18	23	47	97	0.759
7	9	14	38	59	0.751
8	35	30	50	101	0.740
9	36	28	90	140	0.725
10	11	8	24	11	0.706
11	27	27	115	154	0.701
12	25	26	62	65	0.700
13	6	11	24	31	0.692

Continua na próxima página

Tabela 5.3 – continuação da tabela

Aluno	Dias que viu <i>leaderboard</i>	Dias que resolveu exercícios	Exercícios resolvidos antes dos <i>leaderboards</i>	Exercícios resolvidos após os <i>leaderboards</i>	Correlação entre visualização e resolução
14	4	2	61	16	0.689
15	19	21	63	90	0.669
16	10	5	43	22	0.667
17	8	12	40	54	0.643
18	4	5	1	19	0.638
19	23	24	87	80	0.619
20	22	21	21	80	0.618
21	5	2	7	7	0.611
22	15	24	39	106	0.610
23	44	28	57	99	0.600
24	34	27	115	169	0.594
25	10	13	56	53	0.593
26	12	14	104	36	0.589
27	18	19	39	66	0.589
28	13	20	30	73	0.587
29	9	12	26	52	0.575
30	13	16	38	55	0.567
31	18	20	79	103	0.561
32	11	14	70	27	0.557
33	29	39	42	122	0.556
34	17	31	33	119	0.549
35	33	25	22	107	0.548
36	14	17	21	72	0.548
37	29	30	103	107	0.545
38	5	12	36	22	0.543

Continua na próxima página

Tabela 5.3 – continuação da tabela

Aluno	Dias que viu <i>leaderboard</i>	Dias que resolveu exercícios	Exercícios resolvidos antes dos <i>leaderboards</i>	Exercícios resolvidos após os <i>leaderboards</i>	Correlação entre visualização e resolução
39	45	28	60	112	0.540
40	5	13	37	41	0.534
41	32	26	53	118	0.528
42	24	26	31	196	0.522
43	14	27	48	113	0.515
44	25	12	26	36	0.509
45	2	7	11	21	0.508
46	8	6	31	26	0.497
47	11	18	76	91	0.474
48	10	24	43	122	0.473
49	8	23	35	84	0.472
50	6	5	36	8	0.469
51	31	24	0	164	0.468
52	22	25	44	90	0.467
53	6	19	63	72	0.464
54	8	23	60	62	0.462
55	31	16	62	90	0.446
56	27	30	32	144	0.444
57	13	18	71	65	0.432
58	14	23	20	65	0.430
59	5	8	27	23	0.410
60	21	24	31	85	0.399
61	6	26	27	73	0.393
62	5	13	28	51	0.390
63	5	21	63	96	0.389

Continua na próxima página

Tabela 5.3 – continuação da tabela

Aluno	Dias que viu <i>leaderboard</i>	Dias que resolveu exercícios	Exercícios resolvidos antes dos <i>leaderboards</i>	Exercícios resolvidos após os <i>leaderboards</i>	Correlação entre visualização e resolução
64	7	14	61	41	0.384
65	2	11	73	32	0.383
66	13	16	46	59	0.380
67	16	25	26	79	0.380
68	4	11	41	48	0.377
69	4	4	48	15	0.357
70	6	4	35	18	0.354
71	1	6	105	35	0.342
72	20	18	42	50	0.331
73	27	14	31	53	0.308
74	4	15	49	60	0.298
75	10	32	53	57	0.298
76	3	3	16	3	0.295
77	6	17	18	26	0.278
78	5	23	37	65	0.269
79	4	17	35	58	0.257
80	1	13	24	27	0.234
81	3	28	80	118	0.223
82	2	24	36	93	0.212
83	4	11	31	26	0.199
84	24	15	112	105	0.157
85	1	30	29	68	0.136
86	3	8	61	17	0.115
87	0	16	72	52	0
88	0	4	18	11	0

Continua na próxima página

Tabela 5.3 – continuação da tabela

Aluno	Dias que viu <i>leaderboard</i>	Dias que resolveu exercícios	Exercícios resolvidos antes dos <i>leaderboards</i>	Exercícios resolvidos após os <i>leaderboards</i>	Correlação entre visualização e resolução
89	0	5	8	12	0
90	0	13	26	49	0
91	0	2	7	28	0
92	4	0	6	0	0
93	1	4	33	20	-0.05
94	2	3	41	25	-0.07
95	1	1	12	4	-0.1
96	4	7	35	12	-0.121

Dentre os alunos que não visualizaram *leaderboards* e os que apresentaram correlação negativa ou positiva fraca, a maioria manteve uma quantidade de exercícios resolvidos baixa, visualizando os *leaderboards* com uma frequência ainda menor. Nessa maioria também existem alunos que resolveram exercícios em muitos dias diferentes mas visualizaram os *leaderboards* em poucos desses dias. Os demais alunos apresentaram um comportamento mais heterogêneo, variando entre si com relação à frequência de atividade durante a disciplina.

Nós encontramos alunos que aumentaram consideravelmente a quantidade de exercícios resolvidos após a implantação dos *leaderboards*. Dentre esses alunos, uma boa parcela deles apresentou uma correlação alta entre visualização de *leaderboards* e resolução de exercícios.

## 5.4 Discussões

Nesta seção, discutiremos os resultados encontrados neste estudo, contemplando as análises dos resultados trazidos na seção anterior.

### 5.4.1 Visualização dos *leaderboards* por tipo

Os resultados obtidos indicam que o *leaderboard* geral é o tipo mais interessante para os alunos mais ativos, isto é, que resolveram mais exercícios durante a disciplina. A observação do comportamento da maioria desses alunos ao longo do tempo revelou que eles costumam alternar resolução de exercícios com visualização de *leaderboards*, o que representa um indício de que os *leaderboards* os interessaram. É possível que os alunos nas posições do topo do *leaderboard* geral tenham sido engajados pela visibilidade existente para os alunos nas cinco primeiras posições que, diferente dos demais, são visíveis por todos os alunos. Desse modo, estes alunos podem ter se sentido motivados a resolver exercícios para poderem continuar nas posições do topo e assim continuarem sendo vistos pelos seus colegas. Estudos futuros precisam ser conduzidos para que se compreenda melhor o impacto motivacional existente para os alunos em alcançar as posições do topo ou serem visíveis por todos os seus colegas.

Nossos resultados sugerem que os *leaderboards* menores são uma adição positiva ao design de *leaderboards* para disciplinas introdutórias de programação. Nós percebemos que parte dos alunos que fizeram uma quantidade de exercícios intermediária em relação aos seus colegas visualizaram os *leaderboards* menores com uma frequência maior. Em geral, essa diferença ocorre devido à poucos dias em que estes alunos fizeram mais exercícios e visualizaram mais os *leaderboards* que o habitual, alternando entre essas duas ações. Esse resultado sugere que a utilização de *leaderboards* semanais e diários podem motivar mais alunos, mesmo que de forma momentânea. Uma possível investigação a ser feita futuramente é verificar quantos e quais *leaderboards* utilizar para aumentar os seus efeitos positivos, incentivando os alunos a realizarem mais exercícios por mais tempo.

Embora os *leaderboards* possam ser um recurso interessante para diversos alunos, eles não motivam a todos. Percebemos que uma boa parcela dos alunos que não resolviam muitos exercícios antes dos *leaderboards* continuaram com uma produtividade baixa após a sua implantação. A baixa quantidade de visualizações dos *leaderboards* por parte desses alunos sugere o desinteresse que eles demonstraram em relação aos *leaderboards* e, para esses alunos, outras estratégias precisam ser consideradas.

Durante o estudo, os *leaderboards* semanais e diários eram resetados toda semana e todo dia, respectivamente. Na nossa implementação, os dados referentes à colocação final

dos alunos no final do dia ou da semana eram perdidos toda vez que estes *leaderboards* eram resetados. Salvar as informações referentes aos alunos nas posições mais altas nos *leaderboards* semanal e diário e disponibilizá-las aos alunos, como uma forma de deixar registrado o desempenho deles, é uma forma de motivá-los a competir nos *leaderboards* menores. Desse modo, os alunos que obtiveram altas posições nos *leaderboards* semanais e diários em algum momento poderão ser reconhecidos durante o período.

#### 5.4.2 Visualização dos *leaderboards* e resolução de exercícios no tempo

Neste estudo, observamos que a quantidade de visualizações feitas aos *leaderboards* pelos alunos tende a cair com o passar do tempo. Esta redução pode ter sido proveniente pelo efeito da novidade dos *leaderboards* e pela estrutura da disciplina deste trabalho, que oferece suporte a alunos com diferentes ritmos de aprendizado.

O efeito da novidade é reportado na literatura de gamificação [25] e implica que o interesse em alguma tecnologia, ambiente ou situação tende a reduzir no tempo conforme esta se torna menos novidade para um indivíduo. A alta quantidade de visualizações no início reflete o interesse inicial dos alunos pelos *leaderboards*. Conforme os alunos se acostumam com este elemento de jogo, o interesse neles sofre reduções, o que ocasiona uma menor quantidade de acessos.

A disciplina de programação introdutória deste trabalho adota a filosofia de ensino *Mastery Learning* [6] que, dentre suas características, preza pelo ensino personalizado. Nesta disciplina, os alunos podem avançar em ritmos distintos entre si. Existem casos de alunos que cumprem com toda a ementa da disciplina antes do tempo e, a partir deste momento, passam a se dedicar a realização de um projeto da disciplina. Por terem uma outra atividade para fazer dentro da disciplina, estes alunos abandonam a realização de exercícios e param de visualizar os *leaderboards*. Em disciplinas onde os alunos avançam nos assuntos no mesmo ritmo, os resultados podem ser diferentes, pois os alunos irão avançar na disciplina em um ritmo mais lento e, portanto, passam mais tempo em contato próximo com os exercícios da disciplina.

Nós percebemos que, nos últimos dias em que os dados foram coletados, a quantidade de visualizações dos *leaderboards* e de resolução de exercícios por parte dos alunos sofreu uma queda acentuada. Esta queda é decorrente da véspera do final do período, momento este

em que os alunos são submetidos a avaliações e atividades de outras disciplinas, o que lhes reduz o tempo disponível para dedicação a programação.

É importante ressaltar também que, neste estudo, os *leaderboards* foram ligados durante o período letivo e que, no momento em que foram ligados, todos os exercícios resolvidos anteriormente pelos alunos foram contabilizados na pontuação geral dele. Deste modo, a pontuação dos alunos no *leaderboard* geral era desigual logo no início. Esta diferença inicial pode ter desincentivado parte dos alunos, que se viram em desvantagem instantes depois da competição começar.

Nós vimos que ligar os *leaderboards* durante o período letivo pode incentivar os alunos a fazerem exercícios de conteúdos anteriores, pois estes tem o mesmo valor de exercícios de unidades mais avançadas. Isso pode fazer com que alunos façam exercícios sobre assuntos já dominados por eles, o que pode representar um comportamento indesejável. Nós acreditamos que esse comportamento é mitigado se os *leaderboards* forem ligados no momento em que os alunos começam a resolver exercícios. Nesse cenário, os alunos irão estudando os conteúdos aos poucos e realizando exercícios sobre esses assuntos. Conforme o aluno avança na disciplina, passa a ganhar acesso a mais exercícios novos. Por já ter realizado exercícios sobre assuntos anteriores, ele irá resolver exercícios novos, correspondentes ao conteúdo que ele está estudando naquele momento.

### 5.4.3 Relação entre visualização de *leaderboards* e resolução de exercícios

Os nossos resultados mostram a diversidade que existe quanto ao interesse dos alunos em relação aos *leaderboards*. A observação da correlação de Spearman entre os dias em que eles visualizaram *leaderboards* e resolveram exercícios nos revelou resultados diversos entre os alunos.

Uma parcela dos alunos (5.2%) não visualizou os *leaderboards* nenhuma vez durante os 70 dias em que eles estiveram disponíveis. Esse resultado sugere que existem alunos que demonstraram total desinteresse ou aversão aos *leaderboards*. Esse recurso estava disponível na página principal do sistema de correção automática usado pelos alunos na disciplina. Sendo assim, é pouco provável que o aluno não tenha acessado nenhuma vez por conta

de um total desconhecimento do recurso. Uma parte desses alunos resolveram exercícios frequentemente durante a disciplina, o que indica que, mesmo que não tenham acessado os *leaderboards*, demonstravam engajamento com a resolução de exercícios.

A maioria dos alunos que apresentou uma correlação negativa ou positiva fraca entre visualização de *leaderboards* e resolução de exercícios representa estudantes que resolveram poucos exercícios durante a disciplina, mantendo esse padrão após a implantação dos *leaderboards*, os quais foram visualizados com uma frequência menor. Esse resultado representa um indício de que os *leaderboards* não motivam parte dos estudantes que resolvem poucos exercícios durante a disciplina. Para esses casos, outras intervenções motivacionais precisam ser implementadas e testadas.

Por fim, vimos que alguns alunos aumentaram consideravelmente a quantidade de exercícios resolvidos após a implantação dos *leaderboards*, apresentando também uma correlação alta entre visualização de *leaderboards* e resolução de exercícios. Esse resultado nos dá um indício de que os *leaderboards* correspondem a um recurso que incentiva alunos menos proativos a resolverem mais exercícios em uma disciplina introdutória de programação.

# Capítulo 6

## Conclusões e trabalhos futuros

### 6.1 Conclusões

Neste trabalho, avaliamos e discutimos um design de *leaderboards* para disciplinas de programação introdutória com o objetivo de incentivar os alunos a resolverem mais exercícios de programação. Para isso, levantamos requisitos e elaboramos decisões de design, além de realizarmos um experimento controlado de curta duração e um estudo longitudinal durante um período letivo de uma disciplina de programação para avaliação do design.

Nós observamos que a utilização de *leaderboards* com um critério de pontuação que pontua igualmente exercícios sobre diferentes assuntos incentivam alunos de uma turma de programação introdutória a realizarem mais exercícios. Sobre as características dos exercícios realizados pelos alunos, obtivemos indícios de que o *leaderboard* incentivou a realização de exercícios sobre assuntos mais básicos, sem que tenhamos observado redução na realização de exercícios sobre assuntos mais avançados. O critério de pontuação adotado permite que alunos mais atrasados em relação aos seus colegas possam alcançar altas posições do *leaderboard*, compensando o domínio reduzido com o esforço na realização de exercícios dos assuntos que domina, tornando a competição mais disputada e interessante para um conjunto maior de alunos.

Nós percebemos que, durante o período letivo, a quantidade de visualizações dos *leaderboards* feitas pelos alunos é alta no início e então passa a apresentar quedas, o que indica que o interesse dos alunos pelos *leaderboards* reduz com o passar do tempo. Na véspera de fim do período letivo, momento este em que os alunos são submetidos a uma grande quantidade

de avaliações e atividades de outras disciplinas, a quantidade de visualizações dos *leaderboards* e resolução de exercícios apresenta uma queda acentuada, indicando que os alunos se tornam menos ativos nessa época.

Quanto à utilização de *leaderboards* diferentes quanto ao período de tempo que consideram no cálculo da pontuação, nós vimos que o *leaderboard* geral, que contabiliza todas as atividades resolvidas, foi o mais acessado pelos alunos mais ativos durante a disciplina. Os *leaderboards* menores foram acessados em grande frequência, sendo os mais acessados por parte dos alunos que não estiveram em posições do topo no *leaderboard* geral. Vimos que existiram casos de alunos que, mesmo por pouco tempo, apresentaram aumento no número de visualizações de *leaderboards*, principalmente dos menores, e na quantidade de exercícios resolvidos em um mesmo dia se comparado com dias anteriores, acompanhado de uma alternância entre visualização e resolução. Isso corresponde a um indício de que eles foram motivados pelos *leaderboards*, mesmo que por um breve período de tempo.

Nós observamos que os efeitos motivacionais dos *leaderboards* não são encontrados em todos os alunos, existindo até mesmo alunos que não os acessaram nenhuma vez durante um período letivo. Esse resultado reforça a ideia de que a gamificação em contextos educacionais deve envolver a implantação de mais elementos de jogos (e.g. narrativas, *badges*, níveis) para que pessoas que não são motivadas por competição possam se sentir mais engajadas a resolver exercícios.

Esse trabalho tem como contribuição a proposta e a avaliação de um design de *leaderboards* para disciplinas de programação introdutória. Para que pudéssemos investigar os efeitos dos *leaderboards* com maior precisão, decidimos isolá-lo de outros elementos de jogos e do processo avaliativo da disciplina, para que tivéssemos resultados que fossem decorrentes de sua utilização em particular.

## 6.2 Trabalhos futuros

Como trabalhos futuros, almejamos investigar a utilização de mais tipos de *leaderboards* em disciplinas de programação. Planejamos investigar *leaderboards* criados pelos próprios alunos como os *leaderboards* de amigos, onde alunos competem com um subgrupo de alunos definido por eles, além de *leaderboards* baseados em times, onde os alunos formam grupos

e desafiam outros grupos de alunos.

Também pretendemos estudar o impacto da restrição de visibilidade de informação nos *leaderboards*, que foi implementada neste trabalho mas não foi avaliada. Nós queremos verificar se esta ocultação de informação reduz os impactos negativos da exposição existente nos *leaderboards* onde todos podem ver o desempenho de todos e se esta intervenção tem algum efeito na quantidade de atividade desempenhada pelos alunos e na satisfação deles.

Nos estudos presentes nesta dissertação, realizamos uma análise quantitativa dos exercícios resolvidos pelos alunos, mas não realizamos nenhuma análise relacionada com a qualidade do código enviado por eles. Nós queremos investigar se a existência de *leaderboards* tem efeito na qualidade do código enviado pelos alunos. Além disso, queremos verificar se a utilização de *leaderboards* incentivam os alunos à trapaça. Possíveis indícios de trapaça podem ser encontrados em submissão de código igual ou muito semelhante entre dois ou mais alunos distintos.

Neste trabalho, verificamos os impactos de decisões de design referentes aos *leaderboards* em uma disciplina que adota a filosofia de ensino *Mastery Learning*. Como trabalhos futuros, pretendemos investigar os impactos das decisões de design utilizadas neste trabalho em disciplinas onde todos os alunos seguem o mesmo ritmo de avanço.

Por fim, pretendemos expandir a utilização de elementos presentes em jogos no sistema de correção automática, incorporando *badges*, avatares, narrativas e missões. Nós queremos estudar esses elementos de jogos combinados com os *leaderboards*, para verificarmos se a utilização desses elementos potencializam os efeitos dos *leaderboards* quanto à resolução de exercícios e motivação dos alunos.

# Bibliografia

- [1] Business need to get in the game. <https://www.marketingweek.com/2010/09/23/businesses-need-to-get-in-the-game/>, 2010. Online; acessado em 30 de julho de 2016.
- [2] Learning by Playing: Video Games in the Classroom. <http://www.nytimes.com/2010/09/19/magazine/19video-t.html>, 2010. Online; acessado em 30 de julho de 2016.
- [3] Essential Facts About the Computer and Video Game Industry. <http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf>, 2015. Online; acessado em 28 de julho de 2016.
- [4] Elliott M Avedon and Brian Sutton-Smith. *The study of games*. John Wiley & Sons, 1971.
- [5] Richard Bartle. Hearts, clubs, diamonds, spades: Players who suit muds. *Journal of MUD research*, 1(1):19, 1996.
- [6] Benjamin S Bloom. Learning for mastery. instruction and curriculum. regional education laboratory for the carolinas and virginia, topical papers and reprints, number 1. *Evaluation comment*, 1(2):n2, 1968.
- [7] Ilaria Caponetto, Jeffrey Earp, and Michela Ott. Gamification and education: A literature review. In *ECGBL 2014: Eighth European Conference on Games Based Learning*, pages 50–57, 2014.
- [8] Christopher Cheong, France Cheong, and Justin Filippou. Quick quiz: A gamified approach for enhancing learning. In *PACIS*, page 206, 2013.

- [9] Katheryn R Christy and Jesse Fox. Leaderboards in a virtual classroom: A test of stereotype threat and social comparison explanations for women's math performance. *Computers & Education*, 78:66–77, 2014.
- [10] Thomas M Connolly, Elizabeth A Boyle, Ewan MacArthur, Thomas Hainey, and James M Boyle. A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, 59(2):661–686, 2012.
- [11] Eliane Cristina de Araujo, Matheus Gaudencio, Andrey Menezes, Iury Ferreira, Iara Ribeiro, Alberto Fagner, Lesandro Ponciano, Fábio Morais, Dalton S Guerrero, and Jorge A Figueiredo. O papel do hábito de estudo no desempenho do aluno de programação. *Workshop de Educação em Computação, Congresso anual da SBC 2013, Maceió, Brasil. SBC.*, 2013.
- [12] Edward L Deci, Richard Koestner, and Richard M Ryan. Extrinsic rewards and intrinsic motivation in education: Reconsidered once again. *Review of educational research*, 71(1):1–27, 2001.
- [13] Edward L Deci and Richard M Ryan. The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry*, 11(4):227–268, 2000.
- [14] Sebastian Deterding. Gamification: designing for motivation. *interactions*, 19(4):14–17, 2012.
- [15] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. From game design elements to gamefulness: defining gamification. In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, pages 9–15. ACM, 2011.
- [16] Darina Dicheva and Christo Dichev. Gamification in education: Where are we in 2015? In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education*, volume 2015, pages 1445–1454, 2015.
- [17] Darina Dicheva, Christo Dichev, Gennady Agre, and Galia Angelova. Gamification in

- education: a systematic mapping study. *Educational Technology & Society*, 18(3):1–14, 2015.
- [18] Adrián Domínguez, Joseba Saenz-De-Navarrete, Luis De-Marcos, Luis Fernández-Sanz, Carmen Pagés, and José-Javier Martínez-Herráiz. Gamifying learning experiences: Practical implications and outcomes. *Computers & Education*, 63:380–392, 2013.
- [19] K Anders Ericsson, Ralf T Krampe, and Clemens Tesch-Römer. The role of deliberate practice in the acquisition of expert performance. *Psychological review*, 100(3):363, 1993.
- [20] Leon Festinger. A theory of social comparison processes. *Human relations*, 7(2):117–140, 1954.
- [21] Márta Fülöp. *Japanese students' perception of the role of competition in their country*. 1999.
- [22] Stephen M Garcia, Avishalom Tor, and Richard Gonzalez. Ranks and rivals: A theory of competition. *Personality and Social Psychology Bulletin*, 32(7):970–982, 2006.
- [23] Matheus Gaudencio and Dalton S Guerrero. Prog1box: Um ambiente linux sem distratores para o aprendizado de programação. In *Anais do Simpósio Brasileiro de Informática na Educação*, volume 25, page 906, 2014.
- [24] James Paul Gee. What video games have to teach us about learning and literacy. *Computers in Entertainment (CIE)*, 1(1):20–20, 2003.
- [25] Juho Hamari, Jonna Koivisto, and Harri Sarsa. Does gamification work?—a literature review of empirical studies on gamification. In *System Sciences (HICSS), 2014 47th Hawaii International Conference on*, pages 3025–3034. IEEE, 2014.
- [26] Michael D Hanus and Jesse Fox. Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, 80:152–161, 2015.

- [27] Robin Hunicke, Marc LeBlanc, and Robert Zubek. Mda: A formal approach to game design and game research. In *Proceedings of the AAAI Workshop on Challenges in Game AI*, volume 4, page 1, 2004.
- [28] Alexandru Iosup and Dick Epema. An experience report on using gamification in technical higher education. In *Proceedings of the 45th ACM technical symposium on Computer science education*, pages 27–32. ACM, 2014.
- [29] David W Johnson, Geoffrey Maruyama, Roger Johnson, Deborah Nelson, and Linda Skon. Effects of cooperative, competitive, and individualistic goal structures on achievement: A meta-analysis. *Psychological bulletin*, 89(1):47, 1981.
- [30] Jesper Juul. The game, the player, the world: Looking for a heart of gameness. *PLURAIIS-Revista Multidisciplinar*, 1(2), 2010.
- [31] Karl M Kapp. *The gamification of learning and instruction: game-based methods and strategies for training and education*. John Wiley & Sons, 2012.
- [32] Fred S Keller. “good-bye, teacher...” 1. *Journal of applied behavior analysis*, 1(1):79–89, 1968.
- [33] Alfie Kohn. *No contest: The case against competition*. Houghton Mifflin Harcourt, 1992.
- [34] Alfie Kohn. *Punished by rewards: The trouble with gold stars, incentive plans, A’s, praise, and other bribes*. Houghton Mifflin Harcourt, 1999.
- [35] Richard N Landers, Kristina N Bauer, and Rachel C Callan. Gamification of task performance with leaderboards: A goal setting experiment. *Computers in Human Behavior*, 2015.
- [36] Richard N Landers and Amy K Landers. An empirical test of the theory of gamified learning the effect of leaderboards on time-on-task and academic performance. *Simulation & Gaming*, page 1046878114563662, 2015.
- [37] Joey J Lee and Jessica Hammer. Gamification in education: What, how, why bother? *Academic exchange quarterly*, 15(2):146, 2011.

- [38] Mark R Lepper, David Greene, and Richard E Nisbett. Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and social Psychology*, 28(1):129, 1973.
- [39] Edwin A Locke and Gary P Latham. Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist*, 57(9):705, 2002.
- [40] Thomas W Malone. Toward a theory of intrinsically motivating instruction. *Cognitive science*, 5(4):333–369, 1981.
- [41] Jane McGonigal. *Reality is broken: Why games make us better and how they can change the world*. Penguin, 2011.
- [42] Cristina Ioana Muntean. Raising engagement in e-learning through gamification. In *Proc. 6th International Conference on Virtual Learning ICVL*, pages 323–329, 2011.
- [43] Johnmarshall Reeve and Edward L Deci. Elements of the competitive situation that affect intrinsic motivation. *Personality and Social Psychology Bulletin*, 22:24–33, 1996.
- [44] Ute Ritterfeld, Michael Cody, and Peter Vorderer. *Serious games: Mechanisms and effects*. Routledge, 2009.
- [45] Richard M Ryan and Edward L Deci. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, 55(1):68, 2000.
- [46] Riku Saikkonen, Lauri Malmi, and Ari Korhonen. Fully automatic assessment of programming exercises. In *ACM Sigcse Bulletin*, volume 33, pages 133–136. ACM, 2001.
- [47] Katie Salen and Eric Zimmerman. *Rules of play: Game design fundamentals*. MIT press, 2004.
- [48] Jesse Schell. *The Art of Game Design: A book of lenses*. CRC Press, 2014.
- [49] Katie Seaborn and Deborah I Fels. Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, 74:14–31, 2015.

- [50] John M Tauer and Judith M Harackiewicz. The effects of cooperation and competition on intrinsic motivation and performance. *Journal of personality and social psychology*, 86(6):849, 2004.
- [51] Nicholas Yee. Motivations of play in mmorpgs. 2005.