

Definição e Avaliação de uma Abordagem Gamificada para o Ensino de Teste de Software

Eduardo Henrique Nascimento
Universidade Federal do Rio Grande
do Norte
Natal, RN, Brazil

Roberta Coelho
Universidade Federal do Rio Grande
do Norte
Natal, RN, Brazil

Charles Madeira
Universidade Federal do Rio Grande
do Norte
Natal, RN, Brazil

Kláudio Medeiros
Unifacisa Centro Universitário
Campina Grande, PB, Brazil

Lucas Silva
Universidade Federal do Rio Grande
do Norte
Natal, RN, Brazil

Carlos Oliveira Neto
Universidade Federal do Rio Grande
do Norte
Natal, RN, Brazil

ABSTRACT

Testing is one of the most adopted practices in the software industry to reveal failures early in the development process and thus improve the quality of software in production. Despite their relevance, activities related to software testing are usually not taught and practiced enough in undergrad courses, and they are commonly considered demotivating. In this scenario, gamification has been a widely used approach, consisting of using common game elements in non-game contexts, which can make teaching activities and software testing practices more interesting. Therefore, the present work discusses how gamification was applied to support the Software Testing Discipline offered to undergraduate courses on Software Engineering, Computer Science, and Information Technology. A case study was conducted to evaluate the effects of the proposed gamified approach on students' motivation and learning. In this study, the students had to create JUnit test cases applying domain partitioning criteria using both traditional and gamified approaches. We used Latin Square to perform the study. The results obtained from quantitative and qualitative analyses indicate that (i) there was no significant difference in test coverage when comparing both approaches, (ii) the gamified approach led to higher mutation scores, (iii) and gamification contributed positively to student motivation.

KEYWORDS

Teste de Software, Ensino, Gamificação, Design Thinking, Relato de Experiência

1 INTRODUÇÃO

As atividades de Teste de Software têm sido cada vez mais adotadas pela indústria de software como forma de encontrar falhas e mensurar a qualidade do software produzido. Essa relevância se torna ainda mais evidente quando se observam os valores referentes aos prejuízos provenientes de falhas de software. Em 2018, relatos documentados pela Tricentis Software Fail Watch¹ indicaram um custo estimado de \$1,7 trilhão de dólares provenientes de falhas de software. Esse valor inclui uma variedade de custos associados a falhas de software, como perda de receita, custos de reparo, perda de produtividade, custos de suporte ao cliente, dentre outros.

Apesar de sua importância no desenvolvimento de software, o que se observa frequentemente é que as práticas e técnicas de

Teste de Software não são suficientemente abordadas em cursos de graduação na área de Tecnologia da Informação. Nestes cursos, costuma-se haver uma ênfase consideravelmente maior nas atividades de implementação e projeto de software [3] nos semestres iniciais, sendo os conceitos e práticas de Teste de Software explorados em disciplinas intermediárias ou avançadas do curso [30]. Além de questões curriculares, existem também questões envolvendo motivação e engajamento na realização de atividades de Teste de Software. Os alunos frequentemente enxergam as atividades de teste como sendo tediosas e custosas, pouco práticas e muito teóricas, de modo que os estudantes preferem atividades de caráter mais criativo, como geralmente é o caso das atividades de programação e projeto de software [4].

Os problemas relacionados ao Teste de Software encontrados na formação dos estudantes se refletem no mercado de trabalho. Na indústria, é possível observar uma recorrente desmotivação quanto ao projeto e execução de bons casos de teste [27], o que pode representar uma certa negligência e desatenção quanto a este tipo de atividade, visto que estas atividades costumam ser tediosas e difíceis [13].

Diante de situações nas quais se observa desmotivação e falta de engajamento, uma abordagem que tem sido bastante utilizada é a gamificação, que consiste na utilização de elementos de jogos em contextos que não são jogos [18], com o intuito de promover motivação e melhorias no processo de ensino e aprendizagem [21]. Dentre esses elementos, podem ser utilizados mecanismos de progressão (e.g. pontos de experiência), recompensa (e.g. medalhas), competição (e.g. tabela de líderes) e feedback [34]. No contexto de atividades relacionadas ao Teste de Software, abordagens gamificadas vem sendo propostas com o objetivo de engajar desenvolvedores nas atividades de teste [19], como testes exploratórios, revisão de código e teste de mutação [4].

Neste trabalho, descrevemos nossa experiência na elaboração e avaliação de uma abordagem gamificada para apoiar o processo de ensino e aprendizagem da disciplina de Teste de Software I, ofertada para os cursos de Bacharelado em Engenharia de Software, Bacharelado em Tecnologia da Informação e Bacharelado em Ciência da Computação da Universidade Federal do Rio Grande do Norte (UFRN).

Inicialmente, para avaliar a utilidade e necessidades para a gamificação, utilizamos a abordagem Level Up [7]. Uma vez definidos os elementos, dentre eles, pontos e times, realizamos a avaliação da

¹<https://appdeveloperomagazine.com/the-software-fail-watch-report/>

abordagem proposta. Para avaliarmos a abordagem, a mesma foi aplicada a uma turma da disciplina Teste de Software I. Algumas contribuições deste trabalho são as seguintes:

- Mostramos como se deu a concepção da abordagem gamificada através da utilização de Design Thinking - que foca na análise do problema antes de propor uma solução.
- Realizamos uma avaliação da abordagem gamificada comparando com a abordagem tradicional utilizando a estratégia do Quadrado Latino

Os resultados obtidos a partir das análises quantitativa (i.e., taxa de cobertura e score de mutação) e qualitativa (i.e., questionário aplicado aos envolvidos) indicam que a gamificação colaborou positivamente com a motivação dos alunos, aumentou o score de mutação (indicando que um maior número das faltas injetadas foram identificadas pelos testes elaborados através da abordagem gamificada). Além disto, este estudo indicou que não houve diferença na taxa de cobertura obtida utilizando as duas abordagens.

2 CONCEITOS BÁSICOS

2.1 Jogos Digitais e Gamificação

Na sociedade atual, os jogos digitais têm se tornado um tema relevante, tanto como produto de entretenimento quanto no como tema de pesquisas. Nos jogos digitais, constantemente, os jogadores se deparam com desafios longos, complexos e difíceis, para os quais eles exercem um alto grau de dedicação e empenho, que muitas vezes não exercem em contextos de ensino e aprendizagem.

Na literatura, alguns pesquisadores realizaram esforços para compreender como os jogos digitais conseguem motivar, engajar e divertir as pessoas. De acordo com Malone [23], os jogos atendem as necessidades de desafio, curiosidade e fantasia. Por sua vez, Jane McGonigal [24] afirma que os jogos proporcionam experiências de sucesso e conexão social, desenvolvendo persistência e criatividade dos jogadores. Dessa forma, jogos digitais têm sido utilizados dentro de contextos educacionais [32], onde os professores fazem uso dos jogos para propor desafios interessantes e suas soluções, agindo como um mediador entre o aluno e o conteúdo [21].

Apesar de seus benefícios, os jogos digitais para contextos educacionais costumam demandar muito tempo, recursos e esforços de game design para serem desenvolvidos [21]. Diante disso, por volta dos anos 2010, cresceu na literatura uma alternativa aos jogos completos [18], que utiliza partes de jogos, sendo uma abordagem mais simples e barata, mas ainda eficiente [17] [25], chamada de gamificação.

Segundo Deterding et al. [18], gamificação consiste na utilização de elementos de game design em contextos que não são jogos. No contexto da educação, gamificação pode ser utilizada para promover motivação e engajamento aos estudantes [21] [31], incentivando participação, colaboração, exploração e criatividade [10].

Quanto aos elementos de jogos, segundo Werbach e Hunter [34], estes podem ser divididos em:

- **Dinâmicas:** Representam a interação entre jogador e as mecânicas de jogo. São elementos de nível mais alto;
- **Mecânicas:** Representam elementos que orientam as ações do jogador. São elementos de nível intermediário;

- **Componentes:** Representam elementos concretos utilizados no jogo. São elementos de nível mais baixo.

A Figura 1 ilustra exemplos de cada um destes tipos de elementos definidos por Werbach e Hunter.



Figura 1: Pirâmide de elementos de jogos segundo Werbach e Hunter [34].

Além dos elementos de jogos apontados por Werbach e Hunter [34], um outro framework relacionado com a gamificação é o Octalysis, proposto pelo Yu-Kai Chou [11]. O Octalysis é um framework com um design centrado no humano (no inglês, Human-Centered Design) no qual existem oito ativadores motivacionais, para os quais elementos e técnicas existentes em jogos podem ser utilizados para ativá-los. Os motivadores em questão são:

- (1) Significado épico e chamado (*epic meaning and calling*): Motivador relacionado com o envolvimento em uma causa além do indivíduo ou com o senso de que o indivíduo é um dos poucos que pode exercer uma determinada atividade;
- (2) Desenvolvimento e conquista (*development and accomplishment*): Motivador relacionado com o progresso e desenvolvimento do indivíduo em algo, além do senso de conquista pela obtenção de algo;
- (3) Empoderamento de criatividade e feedback (*empowerment of creativity and feedback*): Motivador relacionado com a criação de algo, expressividade e recebimento de feedback relacionado ao que foi criado;
- (4) Propriedade e posse (*ownership and possession*): Motivador relacionado com a aquisição de recursos ou com o senso de melhorar algo que é de posse do indivíduo;
- (5) Influência social e relacionamentos (*social influence and relatedness*): Motivador relacionado com a conexão com outras pessoas ou de influência sobre elas;
- (6) Escassez e impaciência (*scarcity and impatience*): Motivador relacionado com o desejo de se obter algo escasso ou difícil de se ter, ou que não se tem ou não se pode ter imediatamente;
- (7) Imprevisibilidade e curiosidade (*unpredictability and curiosity*): Motivador relacionado com algo para o qual não se sabe o que irá acontecer, que está sujeito à probabilidade.
- (8) Perda e aversão (*loss and avoidance*): Motivador relacionado com o medo ou desconforto de perder algo ou, ainda, com a chance iminente de perder esse algo.

A gamificação está bastante relacionada com a motivação, visto que essa abordagem busca incentivar as pessoas a realizarem comportamentos de interesse e inibir comportamentos indesejados. A motivação é uma construção psicológica que combina dois elementos principais: a energia para realizar uma ação e o direcionamento dessa energia a alguma atividade [33]. Essa motivação também pode se caracterizar de acordo com a sua origem, podendo ser, segundo a literatura, uma motivação intrínseca ou extrínseca. A motivação intrínseca é aquela proveniente do próprio indivíduo, são mais independentes de fatores externos e costumam ser mais duradouras, enquanto a motivação extrínseca é aquela proveniente de um fator externo, que, em geral, funciona temporariamente.

A literatura explora esses tipos de motivações e como promover uma e/ou a outra. Em contextos educacionais, idealmente, é desejável que os alunos sejam intrinsecamente motivados, visto que estes alunos costumam ser mais persistentes, engajados e apresentam melhor armazenamento de informação [15]. Os elementos de jogos, comuns na gamificação, em geral, representam fatores externos e, portanto, motivadores extrínsecos. Os motivadores externos, principalmente aqueles que são tangíveis e esperados, podem desviar a atenção do indivíduo para si, comprometendo ou até mesmo substituindo a motivação intrínseca daqueles que já eram inerentemente motivados a alguma atividade [22].

2.2 Design Thinking

Design thinking é uma abordagem para resolução de problemas que, antes de propor as soluções para um problema, busca compreendê-lo pela perspectiva das pessoas envolvidas, sendo uma abordagem centrada no humano [8]. Desta maneira, além dos aspectos funcionais, a solução final apresenta outras qualidades, como facilidade de uso, diversão e aceitação do usuário.

O design thinking é norteado pelos seguintes princípios:

- **Empatia:** Deve existir um foco no usuário e no cumprimento de suas demandas, o que deve ser considerado durante todo o processo de desenvolvimento da solução, sendo esta feita de forma incremental;
- **Colaboração:** Deve haver uma cooperação entre as pessoas envolvidas na situação, como pessoas interessadas e pessoas detentoras de conhecimento sobre o problema, que irão fornecer seus pontos de vista e contribuições para a solução final;
- **Experimentação:** Deve haver um processo de experimentação da solução, de modo que ela será desenvolvida e testada iterativamente, com a participação das pessoas envolvidas, que irão fornecer feedback constantemente;
- **Otimismo:** Deve haver um senso constante de otimismo por parte dos envolvidos, tendo em mente que a solução para o problema em questão é possível de ser encontrada.

Além dos princípios, o design thinking também se baseia nos seguintes passos, que devem ser executados de forma iterativa, sendo refinados e ajustados ao longo do processo:

- **Descoberta** - Envolve a compreensão do problema e como este afeta as pessoas envolvidas. Neste passo, acontece coleta de dados através de métodos como entrevistas e questionários para melhor entendimento do cenário;

- **Interpretação** - Envolve a análise e interpretação dos dados coletados, visando entender alguns pontos para que possa ser conduzido a elaboração de ideias;
- **Ideação:** Envolve a proposição de ideias, tendo como base os dados coletados e interpretados, de modo que estas tenham impacto significativo quanto ao problema;
- **Experimentação** - Envolve a aplicação prática das ideias estabelecidas, o que inclui elaborar e experimentar um protótipo, além de coleta de feedback dos envolvidos;
- **Evolução** - Envolve uma análise dos feedbacks coletados com os participantes da etapa de experimentação, quando se observará o que funcionou, o que não funcionou e o que precisa de ajustes.

2.3 Level Up

A abordagem gamificada definida e avaliada no presente trabalho teve como base uma metodologia para concepção de experiências gamificadas educacionais chamada de Level Up [7]. Esta abordagem combina elementos do design thinking, como uma análise profunda do problema e suas razões e desenvolvimento iterativo, bem como elementos do Octalysis [11], como elementos e técnicas de jogos que atendem a aspectos motivacionais.

No Level Up, existem cinco etapas a serem seguidas.

- (1) **Investigação do cenário:** Busca entender o cenário no qual se deseja gamificar, compreendendo qual o problema existente e se gamificação pode ajudar na solução do problema;
- (2) **Definição dos objetivos com gamificação:** Busca entender de forma clara e objetiva o que se pretende obter ou alcançar com a gamificação no cenário, analisando os seus aspectos motivacionais;
- (3) **Seleção dos elementos de jogos:** Busca determinar quais os elementos de jogos a serem utilizados, considerando a justificativa do uso e a interação dentro do modelo de cada elemento;
- (4) **Prototipação da abordagem gamificada:** Busca implementar e executar a abordagem gamificada através de um protótipo, definindo quais métricas serão analisadas;
- (5) **Evolução da abordagem gamificada:** Busca analisar a abordagem gamificada através da utilização do protótipo, verificando se os objetivos foram atingidos, como cada elemento de jogo influenciou na abordagem e o que o público achou de cada elemento.

2.4 Critério De Teste Funcional: Partição em Classes de Equivalência

Como testar consiste em executar um programa para uma dada entrada e comparar a saída com a saída esperada, a atividade de testes se depara com o problema de geralmente existir uma grande quantidade de entradas possíveis para se executar um programa [16][26]. Para lidar com este problema da incompletude dos testes [26], foram propostos critérios de testes que nada mais são do que critérios de escolha: *Quais dos dados de entrada possíveis devo escolher para exercitar o programa de forma a aumentar as chances de revelar as falhas?* Dentre os critérios de teste existentes iremos destacar o critério de testes chamado de partição em classes de equivalência, por ser utilizado no contexto deste trabalho. De acordo com

este critério, o testador deve particionar do domínio da entrada em classes de valores equivalentes de acordo com uma dada propriedade. A estes grupos de valores equivalentes para uma dada propriedade chamamos de classes de equivalência [2].

Após o particionamento da entrada, o testador deve selecionar representantes de cada classe de equivalência. Este critério se baseia na hipótese que valores de uma mesma classe de equivalência tem a mesma chance de revelar falhas. Assim, selecionando elementos de cada classe para realizar os testes, teremos um conjunto mínimo de casos de teste que representem (quando combinados) o domínio inteiro [2]. Esse critério, portanto, parte do pressuposto de que um dado de teste em uma classe de equivalência é tão bom quanto qualquer outro caso na mesma classe [2] - no que tange sua capacidade de revelar falhas.

Um exemplo da aplicação deste critério seria: Na escolha de dados de testes para uma função que recebe um CPF, podemos realizar um teste com um CPF válido e outro com CPF inválido. Ou ainda, para testar uma função que recebe um inteiro e calcula o imposto devido, podemos passar 1 representante dos inteiros positivos, o valor zero e outro representante dos inteiros negativos. Mais detalhes sobre este critério de teste podem ser encontrados em [2].

3 ABORDAGEM PROPOSTA

Para a definição e avaliação da abordagem gamificada proposta neste trabalho, foi utilizada a metodologia para concepção de experiências gamificadas educacionais chamada Level Up [7], que combina conceitos de Design Thinking e do framework Octalysis [11] para concepção de uma solução gamificada. Esta metodologia considera os problemas existentes pela perspectiva de indivíduos envolvidos, bem como elementos de jogos que podem ser utilizados para engajar as pessoas dentro de um contexto.

Os passos seguidos para a concepção da presente abordagem seguindo a metodologia Level Up são descritos a seguir.

3.1 Passo 1: Investigação do Cenário

O objetivo desta etapa foi conhecer o cenário para o qual estava sendo elaborado a abordagem gamificada, bem como os problemas existentes que se buscava solucionar com a gamificação dentro deste cenário. A abordagem gamificada teve como escopo a disciplina de Teste de Software I, ofertada paralelamente para os cursos de bacharelado em Tecnologia da Informação, bacharelado em Engenharia de Software e bacharelado em Ciência da Computação da Universidade Federal do Rio Grande do Norte (UFRN).

Para a definição do problema, foi conduzida uma entrevista semi-estruturada com a docente responsável pela disciplina, que elencou problemas observados em turmas anteriores. Dentre estes problemas, foram destacados os de desmotivação e indisposição para criação de casos de teste e subutilização de critérios de teste ensinados na disciplina. Sendo assim, a abordagem gamificada seria voltada para resolver o problema de desmotivação dos alunos diante de atividades envolvendo criação de casos de testes.

Quanto aos alunos, realizamos uma apresentação sobre o conceito de gamificação para a turma, tanto para introduzi-los ao assunto, quanto para direcioná-los para uma conversa, com o objetivo de, nessa conversa, se obter informações relevantes para nortear

a abordagem gamificada, tais como background, expectativas e motivações dos alunos.

3.2 Passo 2: Definição dos Objetivos de Gamificação

Considerando a investigação do problema realizado na etapa anterior, inicialmente, foi estabelecido que o objetivo da abordagem é o de aumentar a motivação dos alunos em relação às atividades de teste. No entanto, se fez necessário um estreitamento deste objetivo, visto que “atividades de teste” é um conceito amplo, além de que existem muitas técnicas de teste diferentes que poderiam ser exploradas.

Para responder essa questão, a docente da disciplina de Teste de Software foi novamente consultada. Neste momento, a docente destacou a dificuldade dos alunos (das turmas anteriores) no entendimento e utilização do critério funcional chamado de Particionamento em Classes de Equivalência (ou Particionamento do Domínio de Entrada [2]). Por consequência, os alunos criavam casos de teste não efetivos que não adotavam o critério proposto.

Neste momento definiu-se, portanto, que o objetivo da gamificação da abordagem proposta seria de aumentar a motivação dos alunos no momento de criar casos de testes que utilizassem o critério de particionamento em classes de equivalência. Assim, utilizando a gamificação, pretendia-se tornar o desenvolvimento desses testes uma atividade mais divertida e proveitosa, buscando, assim, melhorar a qualidade dos casos de teste desenvolvidos.

3.3 Passo 3: Seleção dos Elementos de Jogos

Nesta etapa, os elementos de jogos a serem utilizados na abordagem gamificada foram definidos. Para a escolha destes elementos, tomou-se como base o framework Octalysis² para análise dos motivadores e escolha de técnicas de jogos, além das dinâmicas, mecânicas e componentes de jogos definidas por Werbach e Hunter [34]. Desta maneira, os elementos de jogos escolhidos foram os seguintes:

3.3.1 Pontos e Barras de Progresso. O elemento “pontos” indica, em geral, o progresso de um indivíduo, permitindo, no caso da abordagem gamificada, que o aluno visualize o seu desenvolvimento, que está associado com a obtenção desses pontos. A pontuação é um elemento que motiva o aluno extrinsecamente e pode promover uma atração inicial para a abordagem gamificada [7]. A obtenção de pontos através da realização das atividades pode gerar um estímulo positivo no estudante [11]. Por fim, o elemento de pontos foi integrado com barras de progresso, permitindo o estudante ter uma visualização melhor de seu progresso.

Na abordagem gamificada, os alunos recebem pontos ao realizarem atividades que estão relacionadas com boas práticas de teste de particionamento de equivalência, como a definição de classes de domínio e cobertura destas classes com relação aos requisitos do programa. Além disso, haveria um feedback da média de pontos da turma em relação à pontuação do aluno, com o intuito de motivar o aluno através da visualização da sua pontuação comparada com o desempenho médio de seus colegas.

3.3.2 Times. O elemento “times” indica que indivíduos irão trabalhar de forma cooperativa em busca de algum objetivo, o que irá

²<https://octalysisgroup.com/framework/>

demandar uma participação ativa de todos os membros da equipe. Este elemento de jogo foi incluído como uma alternativa para um problema recorrente em situações competitivas, que é o baixo desempenho de alguns alunos se comparado com seus colegas. Dessa maneira, caso um aluno não esteja bem na competição no critério individual, os colegas de equipe poderão ajudar este aluno, colaborando assim com um aumento na sua motivação e no desempenho.

3.3.3 Combate. O elemento de jogo “combate” foi incluído em conjunto com o de times, de forma que as equipes formadas pelos alunos se enfrentem em desafios envolvendo particionamento de equivalência. Esta competição pode ser uma maneira de incentivar os alunos através de um estímulo à sua competitividade natural, o que poderá estimular a criatividade desses alunos [29].

O combate é utilizado junto dos elementos de times e pontos, visto que equipes irão se enfrentar através da realização de exercícios de particionamento de equivalência, onde os alunos devem definir as classes de domínio partindo da especificação dos programas. Conforme as classes de domínio são criadas, os times ganham pontos de acordo com a qualidade e cobertura dessas classes.

3.3.4 Premiações. As premiações são recompensas, geralmente físicas, que são utilizadas constantemente para motivar o indivíduo a realizar alguma ação, pois a realização desta ação levará à obtenção da premiação. Premiações são motivadores extrínsecos, que idealmente devem ser trabalhados em conjunto com motivadores intrínsecos, já que os motivadores extrínsecos, apesar de poderem funcionar, normalmente têm uma duração curta.

A premiação foi representada através de balões, distribuídos de acordo com a faixa percentual de acertos em cada uma das questões, com cores diferentes, dependendo da faixa de acerto. Entre 100% e 86%, os alunos ganhavam um balão verde. Entre 85% e 71%, o balão era da cor laranja. Entre 70% e 56%, a cor era amarelo e, entre 55% e 0%, vermelho. Caso o aluno realizasse um novo envio e obtivesse um acerto maior, prevalecia a cor do balão do melhor resultado.

3.3.5 Tentativas. Para cada exercício, o aluno poderia fazer dois envios. Desta forma, ao mesmo tempo em que ele teria uma segunda chance para corrigir um envio com problemas, ele teria que utilizar esse recurso com cautela.

3.4 Ferramentas de Apoio

Para a condução deste estudo, diversas ferramentas foram utilizadas para disponibilização dos exercícios, recebimento das respostas, automatização da correção, dentre outros.

- Plataforma SIGAA: plataforma pela qual os exercícios foram disponibilizados;
- Google Formulários: criação de exercícios, coleta, correção e feedback das respostas;
- Google Apps Script: manuseio e mediação entre as respostas dos exercícios nos formulários e planilha de pontuação;
- Google Sites: centralização do acesso aos exercícios e às planilhas.

4 AVALIAÇÃO

A abordagem gamificada foi avaliada, com o objetivo de comparar essa abordagem com uma abordagem tradicional de ensino de Teste de Software. Para realizar esta avaliação seguimos o Projeto de

Tabela 1: Delineamento do Quadrado Latino

Sujeitos	Exercício 1	Exercício 2
Sujeito 1	Tradicional	Gamificado
Sujeito 2	Gamificado	Tradicional

Experimento do Quadrado Latino [6]. O Quadrado Latino já foi utilizado em outros estudos que visavam avaliar os benefícios de técnicas de Engenharia de Software [1] [5] [28].

Os participantes do experimento foram 36 alunos que compunham a turma da disciplina Teste de Software I, oferecida aos cursos de Bacharelado em Engenharia de Software, Ciência da Computação e Tecnologia da Informação da Universidade Federal do Rio Grande do Norte (UFRN). Esta avaliação foi realizada em meados do semestre letivo. Já haviam sido apresentados aos alunos os conceitos básicos de Teste de Software e a ferramenta JUnit. Os alunos tiveram, então, 2 aulas teóricas sobre o critério funcional de Partição em Classes de Equivalência, após as quais foi realizada a avaliação descrita a seguir.

Em conformidade com o delineamento do Quadrado Latino, estes 36 alunos foram aleatoriamente separados em 18 pares - seguindo a estrutura mostrada na Tabela 2. Como fatores de controle do experimento, existem os participantes e os exercícios elaborados, tendo como tratamentos o tradicional (sem a gamificação) e o gamificado.

Quanto aos exercícios, neste estudo foram utilizados dois cenários de teste. O Exercício 1 estava relacionado com a especificação de uma locadora de veículos, onde o aluno precisava considerar o módulo de multa referente ao atraso do veículo para montar as classes de equivalência. Por sua vez, o Exercício 2 estava relacionado com a especificação de um programa de e-commerce, para o qual o aluno precisava criar classes de equivalência relacionadas às regras do módulo de pagamento deste sistema. Para cada exercício, o aluno deveria criar as classes de equivalência e os casos de teste, utilizando o formulário para responder quais estavam corretos. Os alunos também tinham que desenvolver e executar, através do framework JUnit, os testes de unidade referentes às classes de equivalência.

Quanto à avaliação da abordagem gamificada, foram utilizados dois tipos de avaliação, sendo uma qualitativa e uma quantitativa. Para a avaliação qualitativa, foi aplicado um questionário anônimo com o intuito de verificar a satisfação e motivação dos estudantes. Enquanto isso, a avaliação quantitativa foi a partir da coleta de um conjunto de métricas de teste como detalhado nas questões de pesquisa a seguir.

As questões de pesquisa (QP) nortearam o estudo. Para cada questão de pesquisa, definimos um conjunto de métricas (M) que foram coletadas com o intuito de buscar responder cada questão, conforme detalhado abaixo:

- **QP1: Os elementos de jogos exercem influência sobre a motivação dos alunos na realização das atividades de teste?** M1: Os pontos da escala Likert obtidos pelas respostas dos alunos nas afirmativas relacionadas a cada elemento de jogo utilizado.

Tabela 2: Dias dos encontros da avaliação

05 de Junho	10 de Junho	12 de Junho	17 de Junho
Grupo 1 Tradicional	Grupo 2 Gamificado	Grupo 1 Gamificado	Grupo 2 Tradicional

- **QP2: Há diferença na corretude das classes de equivalência definidas na atividade gamificada e na atividade sem gamificação?** M2: Quantidade de classes corretas definidas por cada grupo.
- **Há diferença na cobertura e na efetividade dos casos de teste definidos utilizando a abordagem gamificada e a tradicional?** M3: Quantidade de falhas encontrados, medido através do score de mutação que é obtido por meio da diferença entre a quantidade de mutações injetados e a quantidade de mutações detectadas (ou mortas) pelos testes unitários e M4: Cobertura dos testes unitários implementados.
- **QP4: Como os alunos se posicionaram em relação à abordagem gamificada e tradicional, no que diz respeito à satisfação?** M5: Os pontos da escala Likert obtidos através das respostas dos alunos nas afirmativas relacionadas aos aspectos de satisfação referentes ao: divertimento na abordagem, feedback da correção, ambiente utilizado para a realização dos exercícios.

A avaliação da abordagem foi realizada em quatro dias, seguindo a ordem que foi previamente planejada no Quadrado Latino, de modo que, em cada dia, foi realizada uma atividade diferente com um grupo diferente, seguindo a distribuição detalhada na Tabela 2.

Nos dias em que a abordagem tradicional foi utilizada, a tarefa era disponibilizada no SIGAA. A tarefa continha a especificação do programa, assim como um arquivo .jar, que deveria ser utilizado pelos alunos para executar os testes de unidade que foram desenvolvidos por eles. O envio de solução dos alunos incluiu: as classes de equivalência definidas, a lista dos casos de testes que foram criados e a classe JUnit com os testes implementados e que seriam corrigidos.

Nos dias em que a abordagem gamificada foi utilizada, a tarefa também era disponibilizada no SIGAA e os alunos respondiam em duplas. A tarefa no sistema continha o arquivo .jar com a implementação do programa e o link para o formulário com a descrição dos problemas e perguntas (sobre classes de equivalência e casos de teste) e a lista de checkboxes com as alternativas de respostas. Ao fim da tarefa, os alunos também precisavam enviar via SIGAA a classe de testes JUnit com os testes automatizados.

A avaliação da abordagem foi realizada em quatro dias de aula, seguindo a ordem que foi previamente planejada no Quadrado Latino, de modo que, em cada dia, foi realizada uma atividade diferente com um grupo diferente.

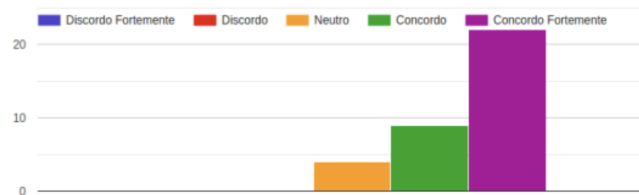
Dentre os 36 alunos que participaram do experimento, 35 responderam um questionário avaliativo. O questionário envolvia dez afirmativas, para as quais foi utilizada a escala Likert, indo de 1 (discordo fortemente) até 5 (concordo fortemente), além de duas questões descritivas baseadas em estudos de Fraser [19]. O questionário pode

**Figura 2: Quanto ao trabalho emde Aplicação da Atividade Gamificada.**

ser acessado no link <https://forms.gle/r83u3UiVMD1Uv3W48>. As respostas do questionário foram anonimizadas e podem ser acessadas no link <https://11nk.dev/z37Mi>.

4.1 QP 1: Os elementos de jogos exercem influência sobre a motivação dos alunos na realização das atividades de teste?

Para a QP1, foram elaboradas questões sobre cada elemento de jogo da abordagem gamificada. Na Tabela 3, estão os motivadores de cada item, bem como os elementos de jogos e os itens do questionário associados com esse motivador. Quanto ao trabalho em times, a grande maioria dos alunos concorda em algum grau que gostaram da experiência de cooperação - ver Figura 3.

**Figura 3: Resposta dos participantes para a afirmativa A1: Eu gostei do trabalho e cooperação em equipe.**

Na atividade, a cor do balão mudava de acordo com a quantidade de alternativas corretas e incorretas respondidas. Os alunos foram questionados (através do survey) sobre a utilização desses balões. As Figuras 4, 5 e 6 apresentam a opinião dos alunos quanto a utilização dos balões.

Considerando as respostas para a afirmativa ilustrada na Figura 5, vimos que as equipes, em sua maioria, discordaram ou foram neutras quanto a se sentirem desestimulados ao observarem o ganho de balões de outras equipes.

Por fim, considerando a afirmativa apresentada na Figura 6, vimos que a maioria dos alunos concordou que a obtenção de balões por parte de outras equipes os motivou a conseguir mais balões.

Resposta a QP1: A maioria dos alunos se sentiu motivada através do ganho de balões, não se sentiu frustrada com o ganho de balões de outras equipes e se sentiu incentivada a responder as questões ao verem que outras equipes estavam ganhando os balões.

Tabela 3: Afirmações a serem avaliadas usando a escala Likert no formulário que buscou responder à QP1.

Ativador Motivacional	Elemento do Jogo	Quantidade de Afirmações	Justificativa
Realização	Pontos	3 (A4, A5, A6)	Indicação de progresso visualização text do desenvolvimento por meio do ganho.
Empoderamento	Premiações	1 (A3)	Recompensar fisicamente para gerar empenho por meio do ganho.
Influência Social	Times	1 (A1)	Incentivo da colaboração entre os membros da equipe.
Escassez	Tentativas	1 (A2)	Tem a característica de recurso limitado, mas visa oferecer uma nova chance para obtenção de sucesso.

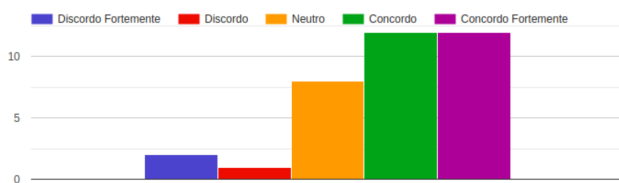


Figura 4: Resposta dos participantes para a afirmativa A4: Me senti estimulado com o ganho dos balões após a conclusão de cada questão.

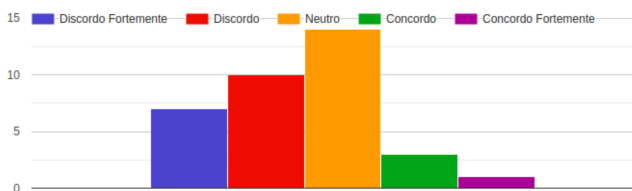


Figura 5: Resposta dos participantes para a afirmativa A5: Me senti frustrado ao ver as outras equipes com mais balões ou com balões melhores que a minha equipe.

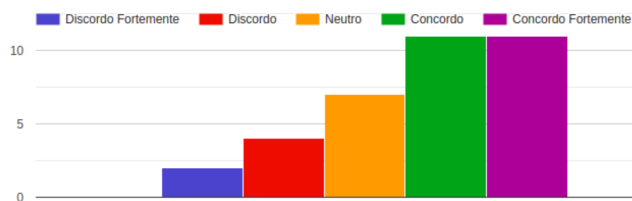


Figura 6: Resposta dos participantes para a afirmativa A6: Ver as outras equipes acumulando balões me incentivou a querer responder mais questões para ganhar mais balões.

4.2 QP2: Há diferença na corretude das classes de equivalência definidas na atividade gamificada e na atividade sem gamificação?

Para a QP2, foram analisadas e corrigidas manualmente todas as classes de equivalência definidas pelos alunos, de modo que cada classe poderia estar incluída em uma dentre três categorias: totalmente correta, parcialmente correta e incorreta.

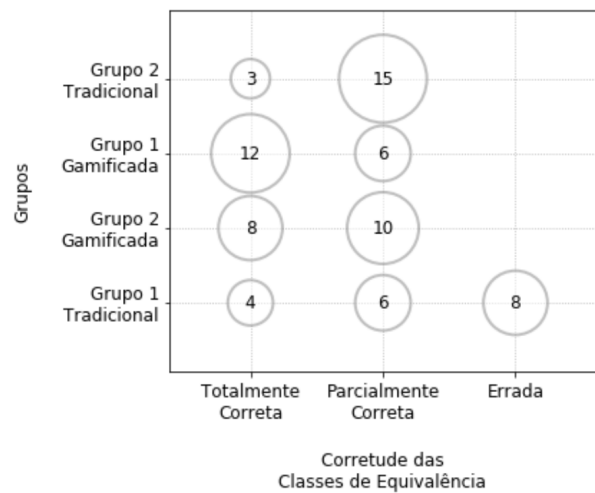


Figura 7: Bubble Plot com a Corretude das Classes de Equivalência por Grupo.

De acordo com o Bubble Plot (ver Figura 7), pode-se perceber que 8 alunos do grupo 1, quando realizaram a atividade com a abordagem tradicional, realizaram submissões erradas. No entanto, este foi o grupo com mais classes de equivalência totalmente corretas quando submetidos à abordagem gamificada, sendo 12 alunos. Os alunos do grupo 2 não apresentaram classes de equivalência erradas, mas foi o grupo que mais submeteu classes que foram classificadas como parcialmente corretas, sendo 15 alunos na abordagem tradicional e 10 na gamificada, sendo também o grupo que menos enviou classes totalmente corretas.

Tabela 4: Abordagens x Corretude das Classes de Equivalência.

Abordagem	Corretas	Parc. Corretas	Erradas
Tradicional (G1+G2)	7	21	8
Gamificada (G1+G2)	20	16	0

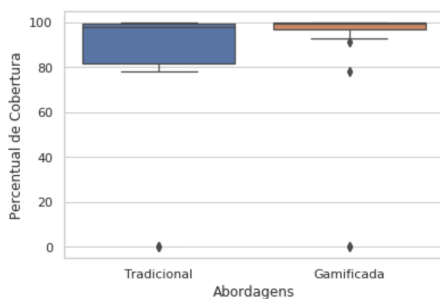
Tabela 5: Mediana e Desvio Padrão do Percentual de Cobertura dos Testes.

Abordagem	Mediana	Desvio Padrão
Tradicional (G1+G2)	97.85	36.15
Gamificada (G1+G2)	99.5	31.09

Resposta a QP2: Atividades submetidas usando a abordagem gamificada apresentaram mais classes de equivalência totalmente e parcialmente corretas quando comparadas com as atividades submetidas utilizando a abordagem tradicional.

4.3 QP3: Há diferença na cobertura e na efetividade dos casos de teste definidos utilizando a abordagem gamificada e a tradicional?

Com relação a QP3, dos casos de teste definidos durante a execução da abordagem gamificada e durante a atividade sem gamificação, foram coletadas as seguintes métricas: i) cobertura de linhas de código dos programas com os testes unitários enviados pelos alunos e ii) o score de mutação para verificar a efetividade dos casos de teste. Para exibição dessas métricas, boxplots foram gerados para mostrar a distribuição do percentual de cobertura dos programas em relação aos testes unitários dos alunos (ver Figuras 8 e 9).

**Figura 8: Box Plot do Percentual de Cobertura dos Testes Unitários.**

Quanto à verificação de efetividade dos casos de teste, para cada sistema, foram introduzidos três defeitos. Os defeitos foram injetados em caminhos distintos do fluxo de controle, de forma que um defeito não tinha influência no outro (e.g. defeitos injetados em métodos distintos que não se comunicavam e defeitos injetados no corpo de blocos if ou else).

Tabela 6: Mediana e Desvio Padrão do Score de Mutação dos Testes Unitários

Abordagem	Mediana	Desvio Padrão
Tradicional (G1+G2)	66.66	38.07
Gamificada (G1+G2)	100.0	39.14

Depois, foi calculado o score de mutação dos testes unitários de cada aluno. Os boxplots com o score de mutação dos testes unitários dos alunos mostram que o score de mutação dos participantes da atividade gamificada é maior que o score de mutação dos participantes da atividade tradicional, visto que o segundo quartil é inexistente e a mediana equivale a 100% (ver Figuras 10 e 11).

Quanto à métrica do score de mutação, também foi realizada a comparação dos resultados obtidos por cada sujeito em ambas as abordagens. Na Figura 10 é apresentada a distribuição dos dados de cada aluno referente aos valores do score de mutação obtidos em cada exercício.

Resposta a QP3: A efetividade dos casos de testes (taxa de mutantes mortos) foi maior nos casos de testes construídos através da abordagem gamificada quando comparados com os gerados na abordagem tradicional. Este estudo não identificou diferença na taxa de cobertura dos casos de testes gerados pelas duas abordagens.

Este resultado corrobora com o que foi apontado no trabalho de Inozemtseva e Reid [20]³: que não existe uma forte correlação entre a taxa de cobertura e a efetividade dos casos de testes.

4.4 QP4: Como os alunos se posicionaram em relação à abordagem gamificada e tradicional, no que diz respeito à satisfação?

Para responder esta questão, foram analisadas as afirmativas do questionário enviado aos alunos após as atividades. Estas afirmativas estão representadas na Tabela 7. Dos 36 alunos, 35 responderam ao questionário. Como podemos observar, a maioria dos alunos: gostou de participar mais da abordagem (94,3%); se divertiu mais durante a abordagem gamificada (94,3%); e preferiu criar classes de equivalência e casos de teste na abordagem gamificada (85,7%).

Resposta a QP4: A maioria dos alunos que participou do estudo gostou mais de construir testes utilizando a abordagem gamificada quando comparada com a tradicional.

4.5 Lições Aprendidas

Através da condução deste trabalho percebemos que não precisamos da implementação de uma ferramenta específica para viabilizar a execução de uma abordagem gamificada e de se obter os benefícios da mesma (e.g. elaboração de quizzes utilizando Google Formulários, apresentação da pontuação utilizando Google Planilhas).

Além disso, a utilização do framework Level Up nos forneceu um passo a passo (já utilizado e avaliado) que nos guiou no entendimento do problema no processo de ensino e aprendizagem e como selecionar elementos de jogos para ajudar na sua resolução.

³Recebeu o prêmio de artigo mais influente em 10 anos no ICSE2024

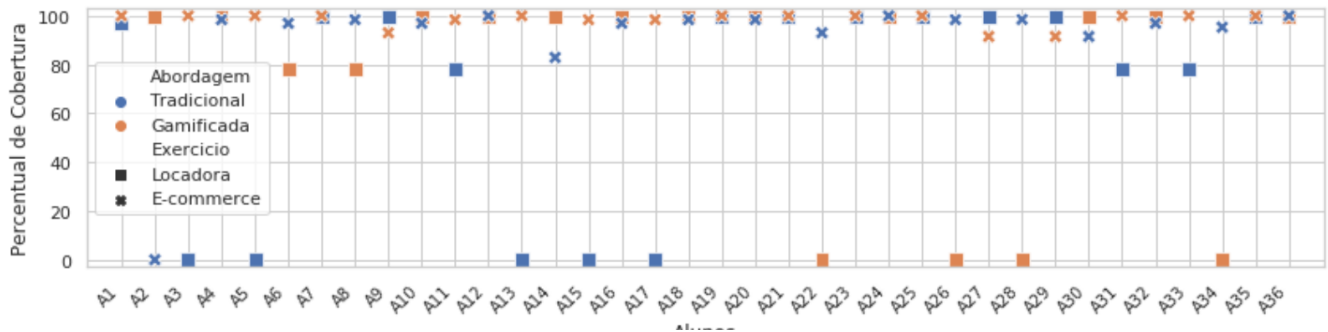


Figura 9: Percentual de cobertura dos testes enviados por aluno.

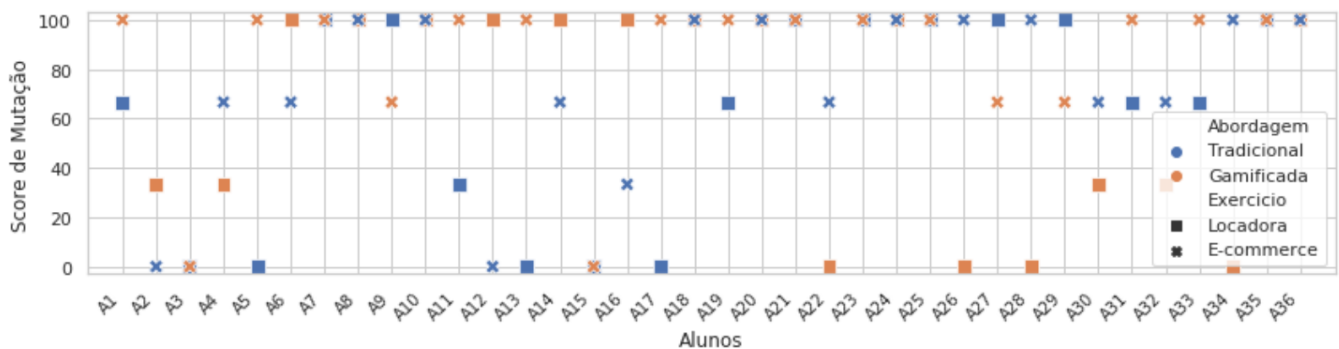


Figura 10: Score de Mutação dos testes enviados por aluno.

Tabela 7: Afirmativas com relação as abordagens gamificada e tradicional e respostas de 35 alunos.

Afirmativa	Gamificada (%)	Tradicional (%)	Nenhuma
Eu gostei de participar mais da abordagem	94,3%	5,7%	0,0%
Me diverti mais durante a abordagem	94,3%	2,85%	2,85%
Eu preferi criar classes de equivalências e casos de testes na abordagem	85,7%	8,6%	5,7%

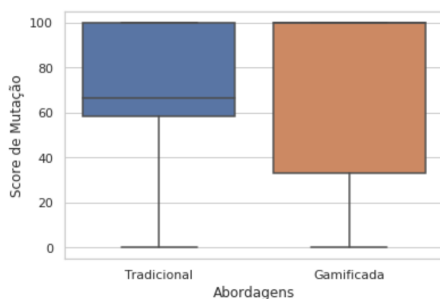


Figura 11: Box Plot do Score de Mutação dos Testes Unitários

Pudemos observar a influência positiva da abordagem gamificada na motivação dos alunos (ver Figuras 3 a 6), além de um aumento na

efetividade dos casos de teste (taxa de mutantes mortos) definidos pelos alunos.

Estes resultados nos levam a crer que precisamos de abordagens de ensino que visem engajar e motivar o aluno, além de simplesmente expor o conteúdo, como é comumente feito em abordagens tradicionais.

4.6 Limitações e Ameaças à Validade

Neste estudo, os alunos da abordagem gamificada possuíam alternativas a serem marcadas, enquanto os alunos da abordagem tradicional não tinham essas alternativas. Esta decisão foi tomada pelo fato da abordagem gamificada utilizar pontos e barras de progresso, e de precisarmos realizar a atualização rápida destes elementos. Assim, a correção automática das alternativas marcadas pelos alunos nos permitia fornecer um feedback rápido para todos os alunos. Alguém pode argumentar que a existência das alternativas pode

ter beneficiado os alunos da abordagem gamificada, uma vez que os alunos da abordagem tradicional não possuíam tais alternativas. Esta ameaça foi mitigada da seguinte forma: (i) para cada questão (por exemplo, quais classes de equivalência podem ser definidas para um dado problema) eram apresentados de 12 a 16 checkboxes com respostas possíveis; (ii) a resposta correta, para uma dada questão, era composta por um conjunto de 3 a 4 checkboxes, que deveriam ser marcados (por exemplo, cada checkbox continha a definição de 1 das classes de equivalência); (iii) como não foi indicado quantos checkboxes formavam a resposta certa, dificilmente o aluno poderia acertar a questão marcando aleatoriamente checkboxes, sem de fato ter realizado a elaboração das classes de equivalência e casos de testes (como na abordagem tradicional).

Outra ameaça à validade neste estudo é o risco dos alunos de uma sessão anterior trocarem informações sobre os exercícios com alunos de uma sessão posterior, fazendo com que alunos tenham algum conhecimento prévio sobre os exercícios. Tal problema foi mitigado alternando os exercícios que eram aplicados e realizando as atividades com cada grupo de alunos em dias diferentes. Além disso, utilizando o Quadrado Latino, buscamos reduzir o efeito de aprendizado. Por fim, os exercícios eram complexos (com várias alternativas a serem selecionadas na abordagem gamificada), o que também dificultou a troca de informações entre eles.

5 TRABALHOS RELACIONADOS

Na literatura, existem outros trabalhos que aplicam gamificação no contexto de Teste de Software. Abaixo, seguem alguns destes trabalhos.

O Code Defenders [29] é um sistema que utiliza da gamificação para criar uma experiência competitiva para ensino e prática de teste de software, mais especificamente, teste de mutação. Nele, os usuários são divididos em dois grupos: o de Attackers (atacantes), que criam mutantes através da injeção de defeitos em um software e o de Defenders (defensores), que devem criar casos de teste que detectam os defeitos. O sistema contém recursos como o de tabela de líderes. Os códigos são escritos em Java e os casos de teste de unidade com o framework JUnit.

O Bug Catcher [9] é um sistema que utiliza elementos de jogos para criar uma competição entre grupos de alunos. Os alunos recebem os requisitos e códigos defeituosos e precisam criar casos de teste para encontrar as falhas nestes programas. Quando um aluno encontra uma falha, é compartilhada com os demais alunos uma tabela. O sistema também possui uma barra de progresso referente a quantidade de falhas encontradas por cada grupo.

Clarke et al. [12] combinaram diversas estratégias de aprendizagem e engajamento em turmas de Teste de Software, como aprendizagem baseada em problemas, aprendizagem colaborativa e gamificação. Como resultados, os autores afirmam que aumentar o tempo dedicado dos alunos a algumas dessas estratégias, dentre elas a gamificação, e diminuir o tempo dedicado a atividades de caráter tradicional, fez com que os alunos tivessem uma melhoria na aprendizagem.

O ambiente gamificado Bug Hunter [14] utiliza elementos de jogos, como avatares, medalhas, pontos e missões para engajar alunos quanto a conceitos de teste e teste funcional. Neste trabalho, os alunos tinham duas rodadas de apresentação de conteúdo seguidas

de um quiz no Kahoot, finalizando com um exercício prático, onde os alunos criavam suites de teste. Os resultados do estudo indicam que a abordagem gamificada foi atrativa, mas que em termos de performance, não se percebeu diferença entre os grupos e, quanto a motivação, os alunos na abordagem tradicional estiveram um pouco mais motivados que os alunos na abordagem gamificada utilizando o Kahoot.

O presente trabalho contrasta com estes principalmente por utilizar a metodologia de gamificação Level Up [7], que parte do Design Thinking, uma abordagem centrada no humano que, antes de propor soluções para um problema, se preocupa em entender os usuários e suas demandas. Nos demais trabalhos, etapas relacionadas com o entendimento do problema pela perspectiva dos envolvidos antes da solução não são mencionadas, ou, quando mencionadas (como em Jesus et al. [14]), não envolvem interação com os alunos como feito no presente trabalho.

6 CONCLUSÕES

Este estudo teve como ponto de partida a desmotivação diante das atividades de teste de software, frequentemente observada em alunos de cursos de tecnologia. Como uma alternativa para lidar com esse problema, existe a gamificação, cuja aplicação tem sido realizada e estudada na literatura, com evidências positivas quanto aos seus efeitos positivos na motivação e engajamento dos estudantes em contextos de teste de software. Diante disso, este trabalho conduziu uma abordagem gamificada para engajar alunos em uma disciplina de Teste de Software a realizarem atividades de teste, focando em atividades envolvendo particionamento de equivalência e derivação de casos de teste.

Para isso, estudos foram realizados com o intuito de compreender a literatura de gamificação aplicada a teste de software, bem como se utilizou uma abordagem para desenvolvimento de experiências gamificadas, que considera demandas e aspectos específicos das pessoas envolvidas no cenário, bem como elementos de jogos que podem ser úteis para motivar essas pessoas dentro desse cenário. Baseado na condução da avaliação da abordagem e nas análises dos resultados obtidos, este trabalho colabora com a literatura indicando que a utilização de gamificação no contexto de disciplinas de teste de software pode aumentar a motivação dos estudantes diante das atividades de teste, e a efetividade dos casos de teste (i.e., score de mutação). Neste trabalho, todavia, não foi observado aumento no percentual de cobertura dos testes elaborados utilizando a abordagem gamificada.

DISPONIBILIDADE DE ARTEFATOS

Mais detalhes sobre a abordagem gamificada avaliada neste artigo podem ser acessados pelo link <https://repositorio.ufrn.br/jspui/handle/123456789/27777>. O questionário aplicado aos alunos que utilizaram as abordagens tradicional e gamificada, as respostas (anonimizadas) e outros artefatos utilizados no estudo estão disponíveis no link: <https://sites.google.com/view/sbeseducationtrack2024>.

AGRADECIMENTOS

Este trabalho é parcialmente financiado pelo INES 2.0 (www.ines.org.br).

REFERÊNCIAS

- [1] Paola Accioly, Paulo Borba, and Rodrigo Bonifácio. 2012. Comparing two black-box testing strategies for software product lines. In *2012 Sixth Brazilian Symposium on Software Components, Architectures and Reuse*. IEEE, 1–10.
- [2] Paul Ammann and Jeff Offutt. 2016. *Introduction to software testing*. Cambridge University Press.
- [3] Tara Astigarraga, Eli M Dow, Christina Lara, Richard Prewitt, and Maria R Ward. 2010. The emerging role of software testing in curricula. In *2010 IEEE Transforming Engineering Education: Creating Interdisciplinary Skills for Complex Global Environments*. IEEE, 1–26.
- [4] Raquel Blanco, Manuel Trinidad, María José Suárez-Cabal, Alejandro Calderón, Mercedes Ruiz, and Javier Tuya. 2023. Can gamification help in software testing education? Findings from an empirical study. *Journal of Systems and Software* 200 (2023), 111647.
- [5] Rodrigo Bonifácio, Paulo Borba, Cristiano Ferraz, and Paola Accioly. 2017. Empirical assessment of two approaches for specifying software product line use case scenarios. *Software & Systems Modeling* 16 (2017), 97–123.
- [6] George EP Box, J Stuart Hunter, and William G Hunter. 2005. Statistics for experimenters: design, discovery and innovation. *Hoboken: Wiley-Interscience* 639 (2005).
- [7] André Luiz de Souza Brito. 2017. *Level Up: uma proposta de processo gamificado para a educação*. Master's thesis. Brasil.
- [8] Tim Brown. 2020. *Design Thinking: uma metodologia poderosa para decretar o fim das velhas ideias*. Alta Books.
- [9] Renée Bryce, Quentin Mayo, Aaron Andrews, Daniel Bokser, Michael Burton, Chelynn Day, Jessica Gonzolez, and Tara Noble. 2013. Bug catcher: A system for software testing competitions. In *Proceeding of the 44th ACM technical symposium on Computer science education*. 513–518.
- [10] Ilaria Caponetto, Jeffrey Earp, and Michela Ott. 2014. Gamification and education: A literature review. In *European conference on games based learning*, Vol. 1. Academic Conferences International Limited, 50.
- [11] Yu-kai Chou. 2019. *Actionable gamification: Beyond points, badges, and leaderboards*. Packt Publishing Ltd.
- [12] Peter J Clarke, Debra L Davis, Ingrid A Buckley, Geoff Potvin, Mandayam Thirunarayanan, and Edward L Jones. 2021. Combining learning and engagement strategies in a software testing learning environment. *ACM Transactions on Computing Education (TOCE)* 22, 2 (2021), 1–25.
- [13] Benjamin S Clegg, José Miguel Rojas, and Gordon Fraser. 2017. Teaching software testing concepts using a mutation testing game. In *2017 IEEE/ACM 39th International Conference on Software Engineering: Software Engineering Education and Training Track (ICSE-SEET)*. IEEE, 33–36.
- [14] Gabriela Martins de Jesus, Fabiano Cutigi Ferrari, Leo Natan Paschoal, Simone do Rocio Senger de Souza, Daniel de Paula Porto, and Vinicius Humberto Serapilha Durelli. 2020. Is it worth using gamification on software testing education? An extended experience report in the context of undergraduate students. *Journal of Software Engineering Research and Development* 8 (2020), 6–1.
- [15] EL Deci and RM Ryan. 2010. Intrinsic motivation. The corsini encyclopedia of psychology (pp. 1–2). *Online Library: Wiley* (2010).
- [16] Marcio Delamaro, Mario Jino, and Jose Maldonado. 2013. *Introdução ao teste de software*. Elsevier Brasil.
- [17] Sebastian Deterding. 2012. Gamification: designing for motivation. *interactions* 19, 4 (2012), 14–17.
- [18] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. From game design elements to gamefulness: defining "gamification". In *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*. 9–15.
- [19] Gordon Fraser. 2017. Gamification of software testing. In *2017 IEEE/ACM 12th International Workshop on Automation of Software Testing (AST)*. IEEE, 2–7.
- [20] Laura Inozemtseva and Reid Holmes. 2014. Coverage is not strongly correlated with test suite effectiveness. In *Proceedings of the 36th international conference on software engineering*. 435–445.
- [21] Karl M Kapp. 2012. *The gamification of learning and instruction: game-based methods and strategies for training and education*. John Wiley & Sons.
- [22] Mark R Lepper, David Greene, and Richard E Nisbett. 1973. Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality and social Psychology* 28, 1 (1973), 129.
- [23] Thomas W Malone. 1981. Toward a theory of intrinsically motivating instruction. *Cognitive science* 5, 4 (1981), 333–369.
- [24] Jane McGonigal. 2011. *Reality is broken: Why games make us better and how they can change the world*. Penguin.
- [25] Cristina Ioana Muntean. 2011. Raising engagement in e-learning through gamification. In *Proc. 6th international conference on virtual learning ICVL*, Vol. 1. 323–329.
- [26] Glenford J Myers, Corey Sandler, and Tom Badgett. 2011. *The art of software testing*. John Wiley & Sons.
- [27] Bret Pettichord, James Bach, and Cem Kaner. 2013. *Lessons Learned in Software Testing: a Context-Driven Approach*. Wiley.
- [28] Márcio Ribeiro, Paulo Borba, and Christian Kästner. 2014. Feature maintenance with emergent interfaces. In *Proceedings of the 36th International Conference on Software Engineering*. 989–1000.
- [29] José Miguel Rojas and Gordon Fraser. 2016. Code defenders: a mutation testing game. In *2016 IEEE Ninth International Conference on Software Testing, Verification and Validation Workshops (ICSTW)*. IEEE, 162–167.
- [30] Lilian Passos Scatalon, Jeffrey C Carver, Rogério Eduardo Garcia, and Ellen Francine Barbosa. 2019. Software testing in introductory programming courses: A systematic mapping study. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*. 421–427.
- [31] Jorge Simões, Rebeca Diaz Redondo, and Ana Fernández Vilas. 2013. A social gamification framework for a K-6 learning platform. *Computers in human behavior* 29, 2 (2013), 345–353.
- [32] Kurt D Squire. 2008. Video games and education: Designing learning systems for an interactive age. *Educational technology* (2008), 17–26.
- [33] Steffen P Walz and Sebastian Deterding. 2015. *The gameful world: Approaches, issues, applications*. Mit Press.
- [34] Kevin Werbach and Dan Hunter. 2015. *The gamification toolkit: dynamics, mechanics, and components for the win*. University of Pennsylvania Press.