

Gamificação Aplicada à Aprendizagem de Critérios de Teste de *Software*

Jonas C. Souza¹, Simone Borges², Vinicius H. S. Durelli¹

¹Universidade Federal de São João del Rei (UFSJ)
São João del-Rei – MG – Brasil

²Universidade Tecnológica Federal do Paraná - UTFPR
Dois Vizinhos – PR – Brasil

{jonas,durelli}@ufsj.edu.br, simoneborges@utfpr.edu.br

Abstract. *Gamification has been extensively used to support the learning process and keep students engaged. Essentially, gamification is centered on employing game-based mechanics and gamelike features to non-game problems. We evaluate the extent to which gamification can be helpful in teaching graph-based software testing criteria. To this end, we designed and implemented Gamifying Graph Coverage Criteria (GGCC), a tool that presents gamified instructions on graph-based criteria. An experiment involving 20 volunteers has shown that participants exposed to the content through GGCC performed better than participants exposed to the content using traditional classrooms. A later survey confirmed this positive attitude from the participants towards GGCC.*

Keywords: *Gamification; Graph coverage criteria, Software Engineering education*

Resumo. *A gamificação tem sido amplamente utilizada no processo de aprendizagem. Esta utiliza recursos semelhantes a jogos para problemas que não são de jogos. Avaliamos até que ponto a gamificação pode ser útil no ensino de critérios de teste de software baseados em grafos. Para isso, projetamos e implementamos o Gamifying Graph Coverage Criteria (GGCC), uma ferramenta que apresenta instruções gamificadas sobre critérios baseados em grafos. Um experimento envolvendo 20 voluntários mostrou que os participantes expostos ao conteúdo por meio do GGCC tiveram um desempenho melhor do que os participantes expostos ao conteúdo em salas de aula tradicionais, percepção confirmada em pesquisa posterior aplicada participantes em relação ao GGCC.*

Palavras-chave: *Gamificação; Critério de cobertura de grafos, Ensino de engenharia de software*

1. Introdução

Um dos maiores desafios enfrentados na educação continua sendo a criação de estratégias e técnicas inovadoras para auxiliar nos processos de ensino e de aprendizagem com o propósito de tornar os estudos mais fáceis e prazerosos. Recentemente, gamificação vem se destacando como uma técnica promissora para o desenvolvimento de sistemas educacionais engajadores. Gamificação é a utilização de mecânicas, estética e dinâmica de jogos em sistemas que não são jogos, para engajar pessoas, motivar a ação, resolver problemas e

até mesmo promover a aprendizagem [Deterding et al. 2011]. O termo gamificação surgiu em 2010, embora o conceito de gamificação já seja aplicado há algum tempo. Apesar de diversos termos paralelos serem encontrados na literatura, nesta pesquisa o termo gamificação é utilizado para designar a técnica. Em um primeiro momento, as pesquisas sobre gamificação estavam concentradas apenas na sua efetividade. Atualmente, sabe-se que gamificação pode ser uma técnica eficaz na educação, todavia existe também um amplo consenso de que o sucesso da gamificação depende do contexto de uso, o que demanda o desenvolvimento de soluções específicas para cada domínio [Toda et al. 2018].

Outro desafio enfrentado cotidianamente diz respeito aos sistemas de *software*. A sua ubiquidade faz com que falhas ou indisponibilidade de algum desses sistemas normalmente resultem em transtornos e prejuízos financeiros. Um estudo realizado pelo Instituto Nacional de Padrões e Tecnologia relatou um prejuízo aproximado de 59,5 bilhões de dólares por ano, causados por falhas em sistemas de *software* [Tassey 2002]. A correção de defeitos que se propagam até o ambiente de produção de um sistema de *software*, por exemplo, pode chegar a ser cem vezes mais caro do que corrigir este mesmo defeito nas fases iniciais do projeto. O teste de *software* é o principal método utilizado para avaliar o processo de desenvolvimento de *software* [Ammann e Offutt 2017]. Na prática, testes de *software* são uma das principais maneiras de mensurar a qualidade do *software* [Thiry et al. 2011], salientando-se ainda que as práticas e técnicas de teste ainda são pouco aplicadas por empresas de desenvolvimento de *software*.

Apesar de sua profunda importância, pouca ênfase é dada, na área acadêmica, ao ensino de atividades de Engenharia de *Software*. Nos cursos de graduação em Ciências da Computação, os tópicos de Engenharia de *Software* e, principalmente, tópicos relacionados ao teste de *software*, tendem a ser abordados superficialmente e são trabalhados em poucas disciplinas [Von Wangenheim e Silva 2009]. Uma das grandes dificuldades no ensino de teste de *software* é motivar os alunos durante o aprendizado. Diversos fatores dificultam o ensino de conceitos e práticas relacionadas ao teste de *software*. Alguns aspectos podem desestimular no ensino de atividades de teste de *software*, como a falta de atividades práticas, falta de motivação e dificuldades para ensinar as habilidades de elaboração e execução de casos de teste [Von Wangenheim e Silva 2009]. A complexidade dos conceitos relacionados ao teste de *software* pode afetar negativamente o processo de aprendizado. Como exemplo, tem-se a falta de motivação dos testadores para executar bons casos de testes. Portanto, pode-se considerar que grande parte dos profissionais não se encontram devidamente preparados para conduzir atividades de teste de *software* de forma apropriada. De fato, um dos problemas da indústria no desenvolvimento de *software* é a falta de profissionais qualificados para implementar técnicas e boas práticas de teste [Thiry et al. 2011]. Dada a recente importância da condução de atividades que aprimorem a qualidade de sistemas de *software*, a principal motivação da pesquisa descrita neste documento é auxiliar profissionais e estudantes na aquisição e retenção de conhecimentos relacionados ao teste de *software*.

Nesse sentido, este trabalho explora como a gamificação pode ser utilizada para auxiliar e engajar estudantes no aprendizado de teste de *software*, mais especificamente no aprendizado de critérios de testes baseados em cobertura de grafos. A fim de abordar tal questão de pesquisa, o restante desse documento está organizado da seguinte forma: na Seção 2, apresentamos o referencial teórico necessário, na Seção 3 o experimento é

descrito em detalhes e os resultados são discutidos na Seção 4. Resultados associados ao questionário aplicado aos participantes após o experimento podem ser vistos na Seção 5, as considerações finais são apresentadas na Seção 6.

2. Referencial Teórico

2.1. Critérios de Teste de Software

Teste de *software* é o principal método usado para garantir a qualidade de sistemas de *software* [Ammann e Offutt 2017]. Neste contexto, a fim de apoiar e guiar atividades de teste, desenvolvedores utilizam critérios de teste. Assim, conhecimento relacionado a critérios de teste de *software* é fundamental para que desenvolvedores de *software* possam produzir sistemas de *software* de qualidade. Critérios de teste baseados em cobertura de grafos são um exemplo de família de critérios que vem sendo amplamente utilizado.

2.2. Critérios de Teste Baseados em Cobertura de Grafos

Critérios de cobertura são um conjunto de regras para determinar quais partes do sistema de *software* devem ser testadas, reduzindo assim o número de dados de teste. Normalmente, um grafo é representado na forma de diagrama como um conjunto de círculos para os nós, unidos por linhas denominadas como arestas. O propósito é abstrair as instruções de um sistema de *software* para um grafo. Os critérios de cobertura de grafos avaliam um conjunto de testes para um artefato de *software* em termos de como os caminhos correspondente aos casos de teste percorrem pelo grafo do artefato de *software* em teste [Ammann e Offutt 2017].

Para qualquer critério de cobertura baseado em grafo, a ideia é identificar os requisitos de teste em várias estruturas no grafo [Ammann e Offutt 2017]. Essa família de critérios inclui critérios cujos requisitos de teste envolvem cobrir todos os nós bem como todas as arestas de um grafo. Para uma melhor definição de critério de cobertura de grafo, classifica-se inicialmente um conjunto de requisitos de teste (TR). Classificam-se critérios para visitar todos os nós e posteriormente todas as arestas do grafo. De forma geral, um critério de cobertura pode ser definido como:

- Cobertura de grafo: Dado um conjunto TR de requerimentos de teste para um critério C, um conjunto de teste T satisfaz C em um grafo G se, e somente se, para cada requerimento de teste tr em TR, existe pelo menos um caminho p , exercido por um caso de teste t em T tal que p contém tr .

Os requisitos oriundos deste critério podem ter seus valores verdadeiros, se o requisito foi atendido, e falso se o requisito não foi atendido. Por exemplo, na Figura 1, os requisitos de teste para a cobertura dos nós são: $TR = \{visita\ 1, visita\ 2, visita\ 3, visita\ 4, visita\ 5, visita\ 6, visita\ 7\}$. Conforme ilustrado na Figura 1, nós iniciais possuem uma seta adjacente ou no topo e nós finais são definidos por bordas mais espessas.

Para cada nó, verifica-se se este foi visitado ou não. Com isso, a definição formal de cobertura do nó é: Para cada nó alcançável n em G, TR contém o predicado “visita n ”. Nota-se que para cada nó alcançável, o requisito de teste contém o predicado “visita o nó”. Segundo [Ammann e Offutt 2017], embora matematicamente precisa, essa notação é muito complicada para uso prático e, por isso, os autores introduziram uma simplificação nos predicados dos requisitos de teste, mudando a definição para:

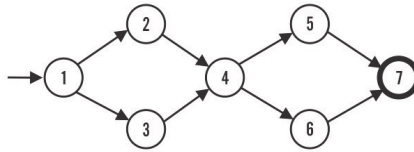


Figura 1. Grafo de entrada e saída única.

- Cobertura de Nós (CN): TR contém todos os nós alcançáveis em G.

Com esta definição, fica entendido que o termo “contém” na verdade significa “contém o predicado visita”. Essa simplificação permite reduzirmos a redação dos requisitos de teste da Figura 1 para apenas contém os nós: $TR = \{1, 2, 3, 4, 5, 6, 7\}$. O caminho do teste $p1 = [1, 2, 4, 5, 7]$ atende aos primeiro, segundo, quarto, quinto e sétimo requisitos de teste e caminho $p2 = [1, 3, 4, 6, 7]$ atende ao primeiro, terceiro, quarto, sexto e sétimo. Portanto, se um conjunto de testes T contiver $\{t1, t2\}$, onde caminho $(t1) = p1$ e caminho $(t2) = p2$, então T satisfaz a cobertura do nó em G. A definição usual de Cobertura de Nós omite a etapa intermediária de identificar explicitamente os requisitos de teste e é frequentemente indicado como abaixo. Observe a economia da definição usada acima em relação à definição padrão.

- Cobertura de Nós (definição padrão): o conjunto de teste T satisfaz a cobertura do nó no grafo G sem, e somente se, para cada nó sintaticamente alcançável n em N, existe algum caminho p no caminho (T) tal que p visita n.

De forma análoga, definimos a cobertura de arestas como a seguir:

- Cobertura de Arestas (CA): TR contém todos os caminhos de tamanho menor ou igual a 1 e que são alcançáveis em G.

2.3. Gamificação

Gamificação consiste no uso de mecânicas, dinâmicas e estética utilizadas no desenvolvimento de jogos, em contextos que não são jogos com o intuito de influenciar a motivação dos usuários em utilizar um sistema ou realizar uma ação [Deterding et al. 2011]. Os efeitos da *gamificação* em contextos educacionais vêm sendo investigados desde seu surgimento, e neste domínio é considerada um processo de planejamento para a adição de elementos de jogos de modo a alterar um dado processo de aprendizagem [Hamari et al. 2014].

Inicialmente, pesquisadores enfatizaram a verificação da eficácia da *gamificação*, entretanto com o avanço das pesquisas, existe atualmente o consenso de que a prática detém o potencial de engajar e motivar, entretanto qualquer solução deve ser cuidadosamente planejada uma vez que a *gamificação* é altamente influenciada pelo contexto onde é aplicada [Toda et al. 2018]. As primeiras soluções de sistemas gamificados faziam uso de elementos de jogos básicos, tais como sistemas de pontuação, troféus, medalhas e moedas virtuais. Este tipo de *gamificação* atualmente é chamado de *gamificação rasa* (*shallow gamification*), cuja principal característica é a ausência de personalização da *gamificação*, ou seja uma solução para todos (*“one size fits all”*). Em contrapartida, o uso de elementos de jogos diferenciados de acordo com o perfil do usuário (perfil de jogador) têm se mostrado capaz de proporcionar soluções mais engajadoras a longo prazo

e é denominada de gamificação profunda (*deep gamification*). Comentar em detalhes as diferenças entre *gamificação* rasa e profunda foge ao escopo deste trabalho. Todavia é importante comentar que ambas as práticas possuem vantagens e desvantagens, não existindo um consenso sobre uma ser inferior a outra, muito pelo contrário, a escolha da prática deve se feita tendo em vista o contexto de uso. Por exemplo, caso um conteúdo a ser aprendido demande um período de aprendizagem longo, recomenda-se técnicas de *gamificação* profunda. Porém, caso a necessidade de engajamento seja por um período mais curto, pode-se utilizar a *gamificação* rasa [Lieberoth 2015]. Além dos benefícios do ponto de vista pedagógico, técnicas de *gamificação* rasa podem demandar menos tempo e custos de desenvolvimento se comparadas à gamificação profunda [Herzig et al. 2015].

3. Metodologia

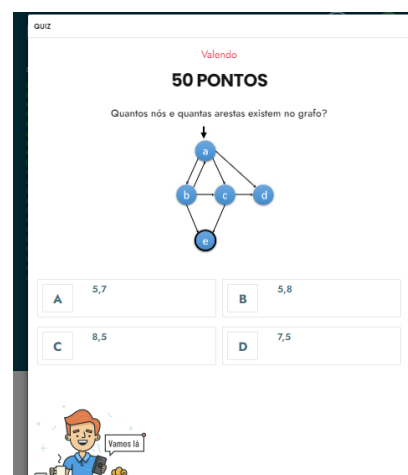
Tradicionalmente, não são oferecidas atividades pós-treinamento para atividades relacionadas ao estudo de critério de cobertura de grafos na engenharia de *software*. Além disso, a falta de diretrizes e de elementos para manter os alunos engajados nos cursos tradicionais torna o processo de aprendizagem menos motivador e mais moroso. Acreditamos que uma abordagem gamificada pode ser empregada para mitigar esse problema. Para investigar os benefícios fornecidos por um ambiente gamificado, foi desenvolvida uma ferramenta chamada *Gamifying Graph Coverage Criteria (GGCC)*.¹

3.1. O GGCC

Por meio de uma plataforma *online*, o GGCC pretende dar suporte a atividades pós-treinamento direcionadas ao ensino de critério de cobertura de grafos. No contexto da ferramenta, essas atividades pós-treinamento seguem uma abordagem de design lúdica, aproveitando os elementos de *gamificação*, como *rankings* e *badges* como forma de motivar o aluno. Um exemplo de elemento lúdico ocorre já na tela de cadastro, na qual usuário pode escolher um *avatar*, como mostrado na Figura 2a abaixo.



(a) Exemplo de tela de cadastro no GGCC



(b) Exemplo de questão aplicada no GGCC

Figura 2. Capturas de tela do GGCC

Por meio do GGCC, o aluno tem contato com o material teórico, apresentado por meio de definições e exemplos, além de realizar *quizzes* sobre o conteúdo apresentado. Ao

¹A ferramenta encontra-se disponível no link: <https://ggcc.dataclix.com.br/>.

acertar questões, o usuário recebe reforços positivos, como mensagens de parabéns, além de efeitos sonoros e visuais, e ao completar um quizz, o GGCC o contempla com a *badge* correspondente.

O conceito de gamificação envolvido no GGCC não tem como finalidade alterar de forma estrutural o processo de aprendizado, mas sim de criar uma motivação extrínseca, e portanto pode ser caracterizado como *shallow gamification*. O uso desta técnica pode ser justificado em função do curto tempo necessário para que o conteúdo em questão seja estudado. Na Figura 2b vemos um exemplo de questão aplicada em um dos *quizzes*, e na Figura 3, vemos um exemplo de *badge* concedida.

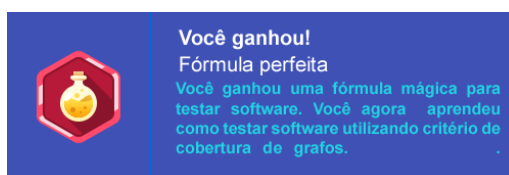


Figura 3. Troféu ao finalizar o último quiz no GGCC

3.2. Protocolo do experimento

Nesse trabalho, exploramos se a *gamificação* pode ter um impacto positivo no reforço pós-treinamento quando comparada com uma abordagem de caráter tradicional. O experimento de que trata esse documento foi desenhado para responder à seguinte questão de pesquisa:

QP: *A gamificação tem de fato um impacto positivo em como estudantes compreendem os critérios de cobertura de grafos nas atividades pós-treino?*

A expectativa era de que os alunos teriam maior taxa de sucesso e satisfação utilizando uma abordagem baseada em jogos para compreender critérios de cobertura de grafos, porque há evidências de que os elementos de *gamificação*, como pontos, rankings e conquistas, transmitem um senso de competência aos alunos e aumentam a sua motivação, melhorando assim o seu desempenho. Postulamos, portanto, que um ambiente gamificado seria mais eficaz em transmitir habilidades e manter os alunos envolvidos, do que uma abordagem mais tradicional para compreender os critérios de cobertura de grafos. Portanto, QP se resumiu em examinar o impacto da *gamificação* no envolvimento dos alunos na compreensão sobre critério de cobertura de grafos a partir da perspectiva de um pesquisador. No contexto deste experimento, utilizamos a seguinte métrica para medir a eficácia da *gamificação*: média de respostas corretas para avaliação de desempenho e média de conquistas.

Nosso experimento foi executado com alunos de Ciências da Computação: alunos de graduação e mestrado da Universidade Federal de São João del-Rei. Tal experimento possuiu apenas um fator, com dois tratamentos distintos. Fator esse definido como a abordagem utilizada para cada um dos participantes: GGCC e TRADICIONAL. Para realização do experimento, os participantes foram divididos em dois grupos, de forma que, ao primeiro grupo, foi apresentado o conteúdo por slide e um questionário com as questões relacionadas. Já para o segundo grupo foi apresentado o slide e o endereço eletrônico do GGCC para utilizarem a ferramenta como atividade de pós-treino. Isto é, o primeiro grupo deveria experimentar o modelo TRADICIONAL e o segundo grupo, o GGCC. Ambos os grupos tiveram contato com os mesmos conteúdos.

PARTICIPANTES	GRUPO	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	FIXAÇÃO	Q21	Q22	Q23	Q24	Q25	TESTE FINAL
1	TRADICIONAL	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	90%	✓	✓	✓	✗	✗	60%
2	TRADICIONAL	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	95%	✓	✓	✓	✓	✓	100%
3	TRADICIONAL	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗	✗	✗	✓	85%	✓	✓	✗	✓	✓	80%
4	TRADICIONAL	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗	✗	✓	80%	✓	✓	✓	✓	✓	100%
5	TRADICIONAL	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	85%	✓	✓	✗	✗	✗	40%
6	TRADICIONAL	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	95%	✓	✓	✓	✗	✗	80%
7	TRADICIONAL	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗	✓	✓	75%	✓	✓	✓	✓	✗	80%
8	TRADICIONAL	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗	✗	✓	75%	✓	✓	✓	✗	✗	40%
9	TRADICIONAL	✓	✓	✓	✗	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗	✗	65%	✗	✗	✗	✗	✗	20%
10	TRADICIONAL	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	95%	✓	✓	✓	✓	✓	80%
11	GGCC	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✗	✓	✗	✓	85%	✓	✗	✓	✓	✗	60%
12	GGCC	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	95%	✗	✓	✓	✗	✗	60%
13	GGCC	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✗	✗	✗	✗	✓	✓	80%	✓	✗	✓	✓	✗	60%
14	GGCC	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	90%	✓	✗	✓	✓	✓	80%
15	GGCC	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	95%	✗	✓	✓	✓	✓	80%
16	GGCC	✓	✗	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	75%	✓	✓	✓	✓	✓	80%
17	GGCC	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	90%	✓	✓	✓	✓	✗	80%
18	GGCC	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	85%	✓	✓	✓	✓	✗	60%
19	GGCC	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	80%	✓	✓	✓	✓	✓	80%
20	GGCC	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	85%	✗	✓	✗	✓	✓	60%

Tabela 1. Tabela com cada nota

4. Resultados e discussão

Na fase inicial do experimento, os participantes foram submetidos a aulas sobre critério de cobertura de grafos. Em seguida, os participantes foram divididos aleatoriamente em dois grupos: TRADICIONAL e GGCC. Logo antes de utilizar a abordagem gamificada, o grupo correspondente foi convidado a uma sessão de treino sobre o GGCC, na qual foram introduzidos todos elementos da ferramenta. Além disso, durante essa sessão, os participantes puderam fazer perguntas sobre a ferramenta. Quanto ao método tradicional, não foi disponibilizada nenhum tipo de assistência.

De acordo com os resultados, a abordagem gamificada foi especialmente efetiva em auxiliar alunos de mais baixo desempenho. As notas mínimas nos quizzes foram de 65% na abordagem tradicional versus 75% na abordagem gamificada. Nas notas finais, a diferença foi ainda mais expressiva: a nota mínima alcançada no teste final foi de 20% na abordagem tradicional, e de 60% na abordagem gamificada.

A comparação dos dois grupos é baseada na comparação da média da nota final dos dois grupos (conforme mostrado na Tabela 1). Dado que foi utilizada uma abordagem Bayesiana para comparação, pymc^3 ² foi utilizado para realizar a inferência que é baseada na distribuição *a posteriori* dos parâmetros. É importante mencionar que a distribuição posterior é obtida por meio da combinação das informações da distribuição *a priori*, i.e., distribuição da probabilidade dos parâmetros antes de se observar a amostra. Conforme mencionado, o principal resultado de uma análise Bayesiana é a distribuição posterior: que consiste basicamente da distribuição dos valores plausíveis considerando os resultados do experimento e o modelo em questão [Gelman et al. 2013]. No contexto do experimento, assumiu-se prioris não-informativas para os parâmetros. Especificamente, adotou-se uma distribuição uniforme ($\theta \sim U(l, h)$ onde l representa o menor valor possível e h o maior) para a distribuição de ambos os grupos: dessa forma, todos os valores plausíveis têm a mesma probabilidade ($l = 0$ e $h = 100$).

Na Figura 4 são apresentados os resultados da distribuição posterior. Conforme pode ser observado, a diferença entre os dois grupos não é substancial. Todavia, de acordo com os resultados do experimento, a distribuição de notas dos participantes do grupo GGCC é mais concentrada em volta da média (*mean*). Adicionalmente, a média do grupo GGCC é ligeiramente maior que a média do grupo tradicional: 70 e 68, respectivamente. O intervalo HPD (*highest-posterior density*) para o grupo tradicional abrange ≈ 50 até ≈ 87 , enquanto o intervalo HDP do grupo GGCC vai de ≈ 62 até ≈ 78 . Apesar do intervalo

²<https://www.pymc.io/welcome.html>.

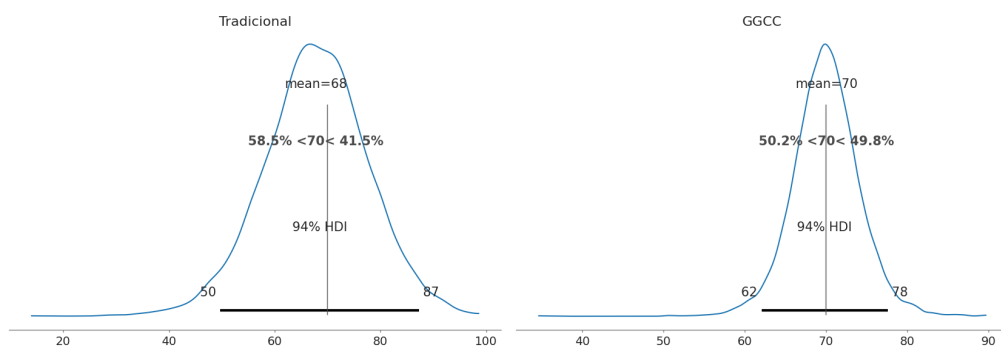


Figura 4. Distribuição posterior com um valor de referência (70) para as notas finais.

HDP do grupo tradicional abranger valores maiores para as notas, uma fração menor do intervalo aparece após a “nota de corte” (i.e., 70): somente 41.5%. O grupo GGCC, por outro lado, mesmo com um intervalo menor, apresenta mais valores após o corte: 49.8%.

5. Questionário Pós-experimento

Esta seção descreve os resultados obtidos por meio do questionário respondido pelos participantes, para responder à seguinte questão de pesquisa: QP: A gamificação tem de fato um impacto positivo em como estudantes compreendem os critérios de cobertura de grafos nas atividades pós-treino? Portanto, o objetivo deste questionário foi avaliar a satisfação, atitude e opinião dos alunos em relação a atividade gamificada pós-treinamento. O questionário³ foi composto por 25 perguntas, divididas em três partes: Q1 a Q4 coletam as informações básicas dos participantes; Q5 a Q9 são perguntas sobre o GGCC (dificuldade com e sem o GGCC); Q10 a Q24 estão relacionadas à experiência dos participantes durante o uso do GGCC. Os participantes do grupo 2 que utilizaram o GGCC foram convidados a responder o questionário ao final do experimento. Apesar de todos os participantes terem respondido o formulário, informamos que seu preenchimento era opcional e anônimo.

5.1. Análise dos resultados

Dez participantes responderam ao questionário, sendo a metade (cinco participantes, ou 50% da amostra) com faixa etária entre 20 e 28 anos, e a outra metade entre 29 e 38 anos. Além disso, sete participantes (70%) afirmaram jogar jogos digitais pelo menos uma vez por semana, dos quais um (10%) afirmou jogar diariamente. Apenas três participantes (30%) afirmaram jogar raramente ou nunca jogaram. Em relação à experiência dos participantes com teste de *software*: seis participantes (60%) afirmaram ter conhecimento apenas com leitura de materiais e/ou cursando a disciplina e três participantes (30%) afirmaram não ter nenhum tipo de conhecimento com teste de *software*. Apenas um participante (10%) afirmou ter conhecimento profissional e atualmente trabalha com teste de *software*.

A Figura 5 apresenta os resultados do questionário sobre as questões Q9 e Q10, que se referem à dificuldade de realizar atividades com e sem o auxílio do GGCC. Os resultados sugerem que os participantes consideraram que as atividades com o GGCC apresentam menos dificuldades em relação às atividades sem o uso do GGCC.

³A íntegra do questionário (no formato PDF) está disponível em <https://ggcc.dataclix.com.br/questionario.pdf>.

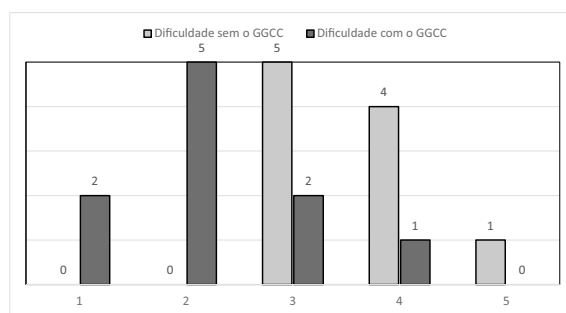
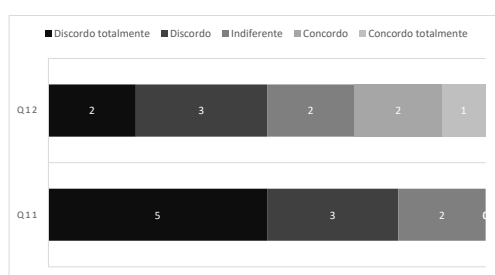
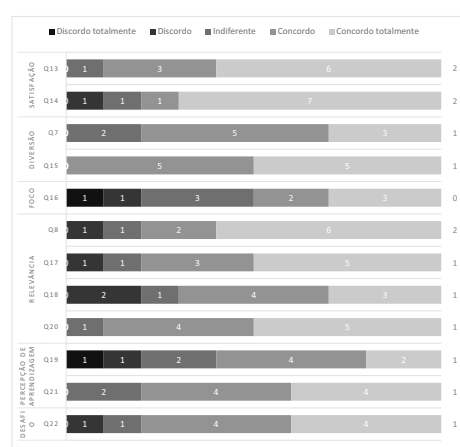


Figura 5. Dificuldade com e sem o auxílio da ferramenta GGCC



(a) Respostas dadas às Questões 11 e 12



(b) Experiência dos participantes

Figura 6

A Figura 6a apresenta as respostas coletadas para as questões Q11 e Q12. Observando as respostas da questão Q11, verifica-se que a maioria dos participantes evitaram pular perguntas. Quanto Q12, apenas um participante (10%) afirmou que tentou responder a todas as perguntas de forma cautelosa e consciente. Portanto, pode-se considerar que alguns participantes tentaram adivinhar a resposta correta para algumas tarefas durante a atividade no GGCC. A Figura 6b apresenta as respostas relacionadas às experiências dos participantes usando o GGCC. As questões foram agrupadas conforme os seguintes fatores: desafio, satisfação, diversão, foco, relevância e percepção de aprendizagem. Nenhuma questão apresentou um valor mediano negativo conforme a escala estabelecida variando de 'Discordo totalmente' (-2) a 'Concordo totalmente' (2). O fator 'foco' apresentou valor mediano de 0, significando indiferença ou opinião mesclada pelos participantes. Para todos os outros fatores, os valores medianos foram positivos. As questões Q13 e Q14 tiveram as respostas 'Concordo totalmente' com mais frequência. Essas questões estão relacionadas à satisfação em utilizar o GGCC como método de aprendizagem, e se o participante indicaria o GGCC para outros alunos que gostariam de aprender e reforçar os conhecimentos em critérios de cobertura de grafos. O aspecto positivo também se deve à questão Q8, sendo a maioria respondendo 'Concordo totalmente'. Este fator considerou a relevância do conteúdo e a abordagem utilizada no GGCC. Todos outros fatores analisados

foram respondidos 'Concordo'. As questões com maior número de respostas positivas foram: Q13 (se o participante ficou satisfeito e realizado ao completar as tarefas do GGCC), Q14 (se o participante recomendaria o GGCC para outros alunos) e Q8 (sobre a abordagem do conteúdo utilizada no GGCC) com 60% de respostas positivas na Q13 e Q14 com 70% de respostas positivas. Q8 ficou com 60% de respostas positivas.

6. Considerações finais

Avaliamos empiricamente se a *gamificação* pode ter um impacto positivo nas atividades pós-treinamento para testar *software* utilizando critério de cobertura de grafos e seus conceitos. Para avaliar a eficácia da *gamificação* neste contexto, usamos a análise bayesiana. Interpretamos esses resultados como suporte para nossa hipótese de que a *gamificação* pode ser aplicada para engajar alunos nas atividades de teste de *software* que são pouco motivadas no ensino. De acordo com os resultados de nosso estudo, em média, mais alunos conseguiram superar a nota de corte de 70% usando a abordagem *gamificada*, quando comparados com os que utilizaram a abordagem tradicional. Esse resultado corrobora nossa hipótese de que o GGCC pode ser uma ferramenta importante no suporte ao ensino de critérios de cobertura de grafos. A satisfação do usuário também foi avaliada e se mostrou bastante positiva.

Referências

- Ammann, P. e Offutt, J. (2017). *Introduction to Software Testing*. Cambridge University Press, 2th edition.
- Deterding, S., Dixon, D., Khaled, R., e Nacke, L. (2011). From game design elements to gamefulness: Defining gamification. In *International Academic MindTrek Conference: Envisioning Future Media Environments*, pages 9–15.
- Gelman, A., Carlin, J. B., Stern, H. S., e Rubin, D. B. (2013). *Bayesian Data Analysis*. Chapman and Hall/CRC, 3 edition.
- Hamari, J., Koivisto, J., e Sarsa, H. (2014). Does Gamification Work? – A Literature Review of Empirical Studies on Gamification. In *2014 47th Hawaii International Conference on System Sciences*, pages 3025–3034. IEEE.
- Herzig, P., Ameling, M., Wolf, B., e Schill, A. (2015). *Implementing Gamification: Requirements and Gamification Platforms*, pages 431–450. Springer.
- Lieberoth, A. (2015). Shallow gamification: Testing psychological effects of framing an activity as a game. *Games and Culture*, 10(3):229–248.
- Tassey, G. (2002). *The economic impacts of inadequate infrastructure for software testing*. National Institute of Standards and Technology, 2002.
- Thiry, M., Zoucas, A., e da Silva, A. C. (2011). *Empirical study upon software testing learning with support from educational game*.
- Toda, A. M., Valle, P. H. D., e Isotani, S. (2018). The dark side of gamification: An overview of negative effects of gamification in education. pages 143–156. Springer.
- Von Wangenheim, C. G. e Silva, D. A. (2009). *Qual conhecimento de engenharia de software é importante para um profissional de software?*, volume 2.