

# Projeto de um *Framework* Conceitual para Avaliação do Desempenho do Aluno em Jogos Sérios

Rhálfeff Nascimento Rodrigues de Oliveira, Rafaela Vilela da Rocha

Pós-Graduação em Ciência da Computação - Universidade Federal do ABC (UFABC)

CEP: 09210-580 – Santo André – SP – Brasil

{rhaleff.nascimento; rafaela.rocha}@ufabc.edu.br

**Resumo:** A avaliação do desempenho do aluno em jogos sérios é uma tarefa complexa, mas necessária para medir a efetividade de tais jogos. Atualmente, muitos jogos sérios são criados sem planejar essa avaliação durante sua produção, o que compromete os resultados da aprendizagem do aluno. Este trabalho objetiva apresentar um projeto de um *framework* conceitual com diretrizes para o planejamento e implementação da avaliação do desempenho de alunos em jogos sérios. Os passos para a pesquisa e desenvolvimento da proposta do *framework* são apresentados e discutidos, tal como alguns exemplos e decisões iniciais.

**Abstract:** Assessing student performance in educational games is a complex task but necessary to measure their effectiveness. Currently, serious games are created without planning this assessment in the development phases, which compromises student learning outcomes. This paper aims to describe a project of a conceptual framework with guidelines for the planning and implementation of the assessment of student performance in serious games. The steps for the research and development of this proposal are presented and discussed, as well as some examples and initial decisions.

## 1. Introdução

Jogos sérios são jogos desenvolvidos com um objetivo pedagógico, diferente de um puro entretenimento, e têm sido aplicados na educação, simulação e treinamento [1-3]. Para serem efetivos, eles precisam oferecer, além de outros elementos, uma avaliação eficaz da aprendizagem do jogador [4]. Rocha e Araújo [2] afirmam que em jogos sérios a avaliação pode ter o propósito de avaliar o jogo sério em si (produto desenvolvido) ou o desempenho humano (auxílio à aprendizagem e desenvolvimento do aluno). No entanto, avaliar se os jogadores aprendem o que devem aprender não é uma tarefa fácil, já que o aprendizado é um processo complexo, difícil de avaliar e quantificar [5]. Para isso, Emmerich e Bockholt [12] recomendam que a avaliação deve ser um elemento integrante da produção de um jogo sério, estando presente em todas as suas fases do ciclo de vida. Entretanto, Oliveira et al. [15] identificaram que a maioria dos principais *frameworks* para produção de jogos educacionais não se preocupa com a avaliação da aprendizagem durante o ciclo de vida de produção desses jogos.

A avaliação de desempenho do aluno é importante porque os jogos sérios são produzidos para apoiar a aquisição de conhecimento e/ou desenvolvimento de habilidades, sendo necessária a mensuração dos resultados alcançados, bem como, o fornecimento de *feedbacks* desses resultados aos interessados (alunos, professores, etc.) [6]. Diferentes modelos de níveis de avaliação foram desenvolvidos para garantir a eficácia da avaliação do desempenho, como, por exemplo, os trabalhos de Kirkpatrick [7], Phillips [8] e Hamblin [9]. Todos estes apresentam níveis de avaliação que vão desde a reação e satisfação do aluno, a aprendizagem até a avaliação dos resultados. Alguns autores citam o uso dos quatro níveis de Kirkpatrick para avaliação do desempenho do aluno em jogos sérios [10, 11, 13]. Emmerich e Bockholt [12] citam que o modelo de Kirkpatrick é uma ferramenta valiosa para planejar e projetar a avaliação.

Na literatura, a avaliação em jogos sérios, em sua maioria, é feita no nível 1 (reação) do modelo de Kirkpatrick, buscando coletar a satisfação, motivação e outras reações do aluno [13, 14]. Entretanto, a avaliação da aprendizagem (nível 2), que mede a mudança de atitudes, conhecimentos e/ou habilidades dos participantes como resultado da participação no programa, muitas vezes é negligenciada [4, 10]. Sendo uma das causas a falta de abordagens (processos, modelos, *frameworks*) que concebem e planejem esta avaliação durante o ciclo de vida de produção dos jogos sérios [2, 12].

Nesse contexto, o objetivo principal deste trabalho é apresentar um projeto de pesquisa e desenvolvimento de um *framework* conceitual com diretrizes para o planejamento e implementação da avaliação do desempenho de alunos em jogos sérios. Para isso, será respondida a seguinte questão de pesquisa: *Como, o que e quando a avaliação do desempenho do aluno deve ocorrer dentro do processo de desenvolvimento de um jogo sério?*. Este trabalho está organizado da seguinte maneira: na seção 2 é apresentada a fundamentação teórica com os principais conceitos envolvidos e os trabalhos relacionados. Na seção 3 é descrita a metodologia. Na seção 4, têm-se os resultados esperados e as discussões. Por fim, as considerações finais e trabalhos futuros, na seção 5, seguida das referências.

## 2. Fundamentação Teórica e Trabalhos Relacionados

A avaliação é um método que tem por objetivo estabelecer os resultados desejados para o avaliado, acompanhar e julgar o valor dos resultados obtidos a partir desses padrões desejados, para reforçar, orientar e/ou corrigir o comportamento do avaliado em suas atividades [11]. A avaliação do desempenho humano determina o grau em que o aluno aplica as competências adquiridas no mundo real, como conhecimentos, habilidades e atitudes [4, 16]. O objetivo principal da avaliação é fornecer *feedback* para os interessados, que pode incluir os alunos, professores/instrutores e coordenadores/gerentes [22].

As avaliações podem ser realizadas em diferentes momentos e ter objetivos específicos. A avaliação pode ser coletada antes (diagnóstica), durante (formativa) ou depois do treinamento (avaliação da reação e do aprendizado) [17]. Além disso, a avaliação somativa visa mensurar a progressão do aluno após um período de aprendizagem [18]. E a autoavaliação tem como objetivo o aluno medir a própria evolução ao final do processo [11].

O modelo proposto por Kirkpatrick [7] é uma estrutura do processo de avaliação do programa de treinamento de quatro níveis: (1) reação, (2) aprendizagem, (3) comportamento e (4) resultados. Os níveis 1 e 2 são baseados nos efeitos imediatos. Na “reação” são avaliados os sentimentos e/ou motivações do usuário após o uso do artefato (atividade) e na “aprendizagem” são identificadas as mudanças no que o usuário conhece sobre o conteúdo após a atividade. Os níveis 3 e 4 são baseados em efeitos de longo prazo. No nível de “comportamento” é avaliado se houve mudanças em relação ao uso do que o aluno aprendeu, ou seja, se ele é capaz de usar em outras situações e ambientes o conteúdo aprendido. No nível de “resultados” é identificado os efeitos da aprendizagem no mundo real.

O ciclo de vida de algumas metodologias para produção de jogos é dividida, basicamente, em *pré-produção* (planejamento), *produção* (análise, projeto, implementação, integração e teste) e *pós-produção* (execução e avaliação dos resultados) [11]. Entretanto, algumas metodologias, processos, artefatos e *frameworks* se limitam ao *design* do jogo e não se preocupam com a avaliação do desempenho do jogador [11, 14, 15, 23]. Dessa forma, é importante definir e usar abordagens durante o ciclo de vida da produção do jogo sério para gerar artefatos que compreendam a avaliação do desempenho do aluno de forma eficaz [23].

Oliveira [15] fez uma análise comparativa de *frameworks* para a produção de jogos educacionais. Como resultado, identificou que apenas três, dos treze *frameworks* comparados, se preocupam com a avaliação da aprendizagem (nível 2 do modelo de Kirkpatrick). O *framework* proposto por Killi [19] sugere o uso de pré-teste e pós-teste para avaliar a aprendizagem do aluno. O trabalho de Westera [20] sugere que o planejamento do modelo da avaliação deve estar associado a qualidade, a velocidade e ao estilo do desempenho do aluno, além do uso do *feedback* em pares para melhorar a qualidade do aprendizado. Ibrahim e Jaafar [21] sugerem que o método de avaliação do aluno seja determinado na fase de pré-produção do jogo e recomenda o uso da autoavaliação relacionada aos objetivos de aprendizagem, estes definidos pela Taxonomia de Bloom. No entanto, nos três *frameworks* não há um planejamento da avaliação em outras fases de produção do jogo educacional.

Rocha [11] apresenta uma metodologia na qual há uma preocupação com o planejamento da avaliação do desempenho do aluno nas fases de produção de um jogo sério. Por exemplo, o uso da dimensão de erros humanos, na fase de análise; e o planejamento das avaliações diagnóstica, formativa e somativa na fase de projeto. Entretanto, a proposta se restringe ao tipo de jogo sério criado (sete fases baseadas na taxonomia de Bloom, para treinamento de competências), e precisa ser adaptada para o planejamento da avaliação em outros tipos de jogos e abordagens.

## 3. Projeto e Metodologia

O presente projeto objetiva desenvolver e avaliar um *framework* conceitual para ser usado no planejamento e implementação da avaliação do desempenho do aluno durante o ciclo de vida de produção do jogo sério. Deverão ser desenvolvidas diretrizes com instrumentos e indicadores que cada fase do ciclo de vida de produção de um jogo sério deve adotar, para se ter uma avaliação eficaz do desempenho do aluno. Para isso, a metodologia de pesquisa e desenvolvimento contemplará os seguintes passos:

- **Revisão bibliográfica:** para compreender os principais conceitos que envolvem este trabalho (já realizado);
- **Revisão sistemática:** para analisar as abordagens (modelos, *frameworks*, processos, métodos e afins) usadas durante o processo de produção de jogos sérios - que contemplam o planejamento e desenvolvimento da avaliação da aprendizagem (em andamento);
- **Análise de jogos que contemplam a avaliação da aprendizagem:** para compreender como os jogos sérios já desenvolvidos abordam a avaliação da aprendizagem e do desempenho do aluno (em andamento);
- **Proposta da solução:** desenvolver uma proposta de *framework* conceitual que possibilite o planejamento e desenvolvimento da avaliação durante o ciclo de vida de produção do jogo sério;
- **Validação por especialistas:** solicitar aos especialistas a avaliação da proposta de solução, para identificar problemas e propor melhorias;
- **Melhoria do *framework*:** desenvolver um *framework*, sendo uma versão melhorada e validada por especialistas da solução proposta, como resultado esperado deste projeto ;
- **Estudo de caso:** Usar e avaliar o *framework* para produzir, pelo menos, um jogo sério.

Os instrumentos e indicadores permitem uma medição e avaliação adequada dos resultados recolhidos, e serão classificados dentro dos níveis 1 e 2 estabelecidos pelo modelo de Kirkpatrick [7]. No nível 1 (reação), exemplos de instrumentos são: os questionários de satisfação proposto por Savi [13] e de autoavaliação e reação de Rocha et al. [14]; e de indicadores são: o grau de interesse nas atividades pedagógicas, o grau de satisfação com as fases do jogo, entre outros. Para o nível 2 (aprendizagem), os exemplos de instrumentos são: uso de pré-teste e pós-teste; e como indicadores: o número de acertos das questões, a porcentagem de alunos que obtiveram conceitos "Bom" e "Muito Bom", a duração da resolução das atividades, entre outros.

#### 4. Discussões

Estudos indicam que alguns jogos sérios não obtiveram sucesso por não apresentarem uma avaliação eficaz [12]. Isso corrobora com o fato de que quando a avaliação do desempenho do aluno é planejada e estruturada desde a fase de pré-desenvolvimento do jogo sério, há evidências mais concretas e confiáveis que o aluno foi capaz de adquirir os conhecimentos, habilidades e atitudes desejáveis [5, 12]. O *framework* conceitual que será proposto neste projeto de pesquisa relatado definirá “o que” e “como” será avaliado e “quando” cada uma ocorre. Sabe-se que, para isso, o *framework* deverá se apropriar dos níveis 1 e 2 do modelo de Kirkpatrick [7], integrado com os objetivos de aprendizagem de cada jogo sério, que poderá usar tal *framework* para definir seus instrumentos da avaliação do desempenho do aluno.

O *framework* a ser desenvolvido visa que os jogos sejam projetados considerando a avaliação do desempenho e o *feedback* imediato e constante. Isso acontecerá por meio de recursos que serão selecionados em uma revisão sistemática, tais como, questionários, atividades para pré-teste e pós-testes, e depois estruturados e sistematizados no *framework* proposto. O *framework* será avaliado e validado por especialistas, os quais analisarão a qualidade e completude das diretrizes (instrumentos e indicadores). O principal desafio é desenvolver um *framework* que seja universal e flexível a qualquer tipo de jogo sério que possa ser produzido.

#### 5. Considerações Finais

A avaliação do desempenho de alunos em jogos sérios é uma área com oportunidades de pesquisa e desenvolvimento. Alguns jogos sérios são produzidos sem o planejamento da avaliação do desempenho em seu ciclo de vida. Essa prática torna o jogo sério um instrumento, muitas vezes, ineficiente ou pouco eficiente. Este artigo apresenta um projeto de pesquisa e desenvolvimento de um *framework* conceitual para o planejamento e implementação da avaliação do desempenho do aluno no processo de produção de um jogo sério.

Como trabalhos futuros, pretende-se concluir os passos descritos na seção 3, da revisão sistemática até o estudo de caso com o uso do *framework*. Espera-se que jogos sérios sejam produzidos, usando o *framework* proposto e considerando os aspectos de avaliação do desempenho do aluno.

#### Referências

[1] de Gloria, A.; Bellotti, F.; Berta, R.; Lavagnino, E. (2014). “Serious games for education and training”. *International Journal of Serious Games*, p. 1–15.

- [2] Rocha, R. V.; Araújo, R. B. (2013). “Avaliação de desempenho humano como parte integrada da metodologia de criação de jogos sérios para treinamento”. XXIV Simpósio Brasileiro de Informática na Educação, p. 144–153.
- [3] Zyda, M. (2005). “From visual simulation to virtual reality to games”. *Computer*, v. 38, p. 25–32.
- [4] Salas, E.; Rosen, M. A.; Held, J. D.; Weissmuller, J. J. (2009). “Performance measurement in simulation-based training: a review and best practices”. *Simulation Gaming*, v. 40, p. 33–45.
- [5] Emmerich, K.; Bogacheva, N.; Bockholt, M.; Wendel, V. (2016). Operationalization and Measurement of Evaluation Constructs. In: Dörner R., Göbel S., Kickmeier-Rust M., Masuch M., Zweig K. (eds), “Entertainment Computing and Serious Games”. *Lecture Notes in Computer Science*, v. 9970. Springer, Cham.
- [6] Bellotti, F.; Kapralos, B.; Lee, K.; Moreno-Ger, P.; Berta, R. (2013). “Assessment in and of serious games: an overview”. *Advances in Human-Computer Interaction*. p. 1–11.
- [7] Kirkpatrick, D. L.; Kirkpatrick, J. D. (2006). “Evaluating Training Programs: The Four Levels”. San Francisco, CA, USA: Berrett-Koehler, 3th edition.
- [8] Phillips, P. P. (2005). “Return on investment: (ROI) basics”. American Society for Training and Development: Baltimore.
- [9] Hamblin, A. C. (1978). “Avaliação e controle do treinamento”. São Paulo: McGraw-Hill do Brasil.
- [10] Drummond, D.; Monnier, D.; Tesnière, A.H.A. (2017). “A systematic review of serious games in asthma education”. *Pediatr Allergy Immunol*. v. 28, p. 257-265.
- [11] Rocha, R.V. (2014). “Metodologia iterativa e modelos integradores para desenvolvimento de jogos sérios de treinamento e avaliação de desempenho humano”. Tese (Doutorado em Ciência da Computação) – Departamento de Computação, Universidade de São Carlos, São Carlos.
- [12] Emmerich, K.; Bockholt, M. (2016). Serious Games Evaluation: Processes, Models, and Concepts. In: Dörner R., Göbel S., Kickmeier-Rust M., Masuch M., Zweig K. (eds) “Entertainment Computing and Serious Games”. *Lecture Notes in Computer Science*, v. 9970. Springer, Cham.
- [13] Savi, R.; Gresse Von Wangenheim, C.; Ulbricht, V. R.; Vanzin, T. (2010). “Proposta de um modelo para avaliação de jogos educacionais”. *RENOTE. Revista Novas Tecnologias na Educação*, v. 8, n. 3, p. 1-10.
- [14] Rocha, R.V.; Bittencourt, I.I.; Isotani, S. (2015). “Avaliação de Jogos Sérios: questionário para autoavaliação e avaliação da reação do aprendiz”. In: XIV Simpósio Brasileiro de Jogos e Entretenimento Digital, 2015 - Trilha de Arte & Design, p. 1-10.
- [15] Oliveira, R. N. R.; Cardoso, R. P.; Braga, J. C. B.; Rocha, R. V. (2018). “Frameworks para Desenvolvimento de Jogos Educacionais: uma revisão e comparação de pesquisas recentes”. In: XXIX Simpósio Brasileiro de Informática na Educação, p. 854-863.
- [16] Salas, E.; Salas, M. A. (2007). “Best practices for performance measurement in military simulation-based training: observations from the field”. Report - Department of Psychology, IST/University of Central Florida, p. 33–45.
- [17] Queiroga, F.; Andrade, J. M.; Borges-Ferreira, M. F.; Nogueira, R.; Abbad, G. S. (2012). “Medidas de aprendizagem em TD&E: fundamentos teóricos e metodológicos”. Abbad, G.S.; et al. *Medidas de Avaliação em Treinamento, Desenvolvimento e Educação: ferramentas para gestão de pessoas*. Porto Alegre: Artmed, p. 108–126.
- [18] Perrenoud, P. (1999). “Avaliação: da excelência à regulação das aprendizagens entre duas lógicas”. Porto Alegre: Artes Médicas.
- [19] Kiili, K. (2005). “Digital game-based learning: Towards an experiential gaming model”. *Internet and Higher Education*, v. 8, p. 13-24.
- [20] Westera W.; Nadolski R.; Hummel H.; Wopereis I. (2008). “Serious games for higher education: a framework for reducing design complexity”. *Journal of Computer Assisted Learning*, v. 24, p. 420–432.
- [21] Ibrahim, R.; Jaafar, A. (2009). “Educational games (EG) design framework: combination of game design, pedagogy and content modeling”. In: *International Conference on Electrical Engineering and Informatics*, p. 293-298.
- [22] Zinovieff, M.A. (2008). “Review and Analysis of Training Impact Evaluation Methods, and Proposed Measures to Support a United Nations System Fellowships Evaluation Framework”. Review Report. Department of Human Resources for Health.
- [23] Rocha, R. V.; Araujo, R. B.; (2013). “Metodologia de Design de Jogos Sérios para Treinamento: ciclo de vida de criação, desenvolvimento e produção”. *Proc. of XII Simpósio Brasileiro de Jogos e Entretenimento Digital - SBGames 2013*, pp.1-10, 2013.