

# AvaliaJS: Planejamento da Avaliação do Desempenho de Alunos em Jogos Sérios

Rháleff N. R. Oliveira<sup>1</sup>, Rafaela V. Rocha<sup>1</sup>

<sup>1</sup> Pós-Graduação em Ciência da Computação - UFABC - Santo André - SP - Brasil

{rhaleff.nascimento, rafela.rocha}@ufabc.edu.br

**Abstract.** *Serious Games (SG) should be designed with a focus not just on reaction but on evaluating student performance, which has been neglected. This work aims to develop and evaluate AvaliaJS, a conceptual model (which has two artifacts: canvas model and project document) to structure, guide and support the planning of the conception and execution of the evaluation. The artifacts were applied to three ready-made games as a proof of concept, and the quality of the model was evaluated by six experts in SGs development and evaluation/assessment, using a questionnaire. The results show that the model can be used to support teams in planning, documentation and development of artifacts, data collection and execution of the evaluation and assessment.*

**Resumo.** *Jogos Sérios (JS) devem ser projetados com foco não apenas na reação, mas na avaliação do desempenho do aluno, o que tem sido negligenciado. Este trabalho visa desenvolver e avaliar o AvaliaJS, um modelo conceitual (que possui dois artefatos: modelo canvas e documento de projeto) para estruturar, orientar e apoiar o planejamento da concepção e execução da avaliação. Os artefatos foram aplicados a três jogos já prontos, como prova de conceito. A qualidade foi avaliada por seis especialistas em desenvolvimento e avaliação de JSs, por meio de um questionário. Os resultados apontam que o modelo pode auxiliar na documentação e desenvolvimento dos artefatos, coleta de dados e execução da avaliação.*

## 1. Introdução

Jogos Sérios (JSs) são desenvolvidos com propósitos que vão além do entretenimento, como o ensino e aprendizagem, reabilitação, treinamento e simulação (Abt, 1987). JSs devem ser projetados com objetivos de avaliar o progresso e os resultados do aprendizado, dar *feedback* constante e apropriado ao jogador e personalizar os objetivos de aprendizagem, com foco na avaliação do desempenho do aluno (Rocha, 2014; Trybus, 2010). Entretanto, a literatura aponta que, no contexto de JSs, a avaliação se concentra nos aspectos da reação do usuário em relação ao jogo, como usabilidade, jogabilidade e satisfação (Emmerich e Bockholt, 2016; Oliveira *et al.*, 2019). Ao passo que a avaliação da aprendizagem é negligenciada e muitas vezes não planejada; por falta de abordagens (métodos, metodologias, *frameworks*, modelos) que possibilitam criar e planejar essa avaliação durante o ciclo de vida de produção do JS (Oliveira *et al.*, 2018; Chaudy e Connolly, 2019). Isso pode impactar na qualidade dos JSs produzidos, na sua efetividade e alcance dos objetivos planejados (Rocha *et al.*, 2015).

A avaliação pode ser definida como o processo de julgamento do desempenho do aluno (para reforçar, orientar e corrigir suas competências) ou eficácia do JS (para melhorá-lo, *e.g.*, em aspectos de usabilidade) (Daoudi *et al.*, 2017). A avaliação pode

ser: (1) da *reação* (“os alunos gostaram do jogo?”), (2) da *aprendizagem* (“os alunos aprenderam com o jogo?”), (3) do *comportamento* (“os alunos estão aplicando os novos conhecimentos?”) ou (4) do *resultado* (“o jogo teve impacto nos resultados?”), conforme níveis de avaliação de Kirkpatrick e Kirkpatrick (2016). A avaliação da aprendizagem define o grau de assimilação do conteúdo, transmitido pelo objeto de aprendizagem, *i. e.*, verifica o alcance dos objetivos instrucionais (conhecimentos, habilidades e atitudes) (Kirkpatrick e Kirkpatrick, 2016). A avaliação do desempenho é definida como o processo de coleta, revisão e uso de dados, para fornecer *feedback* ao aluno sobre seus erros e acertos e melhorar seu desempenho atual (Bellotti *et al.*, 2013; Daoudi *et al.*, 2017). O *feedback* é a informação fornecida por um agente sobre aspectos de desempenho ou compreensão e avaliação complementar (Hattie e Timperley, 2007). Em JSs, o *feedback* pode ser usado para relatar o progresso e interação do usuário no jogo, e da sua aprendizagem (Rocha, 2014; Chaudy e Connolly, 2019).

Por se tratar de um tema complexo e abrangente, a avaliação em JSs implica em desafios e dificuldades, que precisam ser solucionados com propostas de estudos direcionados, principalmente, sobre a avaliação da aprendizagem e desempenho (Rocha *et al.*, 2015; Oliveira *et al.*, 2020). Considerando que a maioria das abordagens para avaliação em JSs foca na avaliação da reação do jogador, por meio de questionários e entrevistas aplicados após a interação com o jogo, há a necessidade de desenvolver modelos que auxiliem na coleta e análise de dados para avaliar a aprendizagem e o desempenho, bem como, usá-lo para avaliar e melhorar o próprio jogo (Salas *et al.*, 2009; Oliveira e Rocha, 2020b). Na perspectiva do aluno, é possível desenvolver jogos em que a avaliação apareça de maneira automática e implícita ao jogador, que forneça *feedback* imediato e contínuo, com objetivo de rastrear e coletar informações para melhorar seu desempenho (Rocha *et al.*, 2015; Oliveira *et al.*, 2020). Na perspectiva do professor, a avaliação deve permitir o acesso a relatórios estatísticos de desempenho do aluno, de maneira clara e direta e de fácil manipulação, para análise de dados e aferição dos resultados (Rocha, 2014; Oliveira *et al.*, 2020). Nessa perspectiva, é essencial o planejamento da avaliação definindo “o que”, “como”, “quem”, “por que” e “quando” ela deverá ser executada, integrado aos diferentes pontos de vista das áreas afins de desenvolvimento de JSs (Falchikov, 2005; Rocha *et al.*, 2015).

Este trabalho objetiva propor o AvaliaJS, um modelo conceitual para estruturar, orientar e apoiar a concepção e execução da avaliação em JSs. O AvaliaJS possui um modelo canvas, para planejamento em alto nível da avaliação, e um documento de projeto, para descrição em baixo nível do planejamento realizado no canvas. Os artefatos foram avaliados por meio de: (1) prova de conceito: para exemplificar e analisar a cobertura dos artefatos propostos; e (2) painel de especialistas: para avaliação técnica do modelo e coleta de recomendações, por meio de um questionário.

## 2. Trabalhos relacionados

Para compreender o estado da arte, foram analisados 27 estudos sobre avaliação em JSs, divididos em três grupos: (1) *abordagens para desenvolvimento de jogos e JSs*; (2) *modelo canvas para design de jogos em geral*; e (3) mapeamentos e revisões sistemáticas da literatura *que analisam a avaliação em JSs*. As referências dos trabalhos analisados podem ser encontradas em [bit.ly/RefTrabRelac](https://bit.ly/RefTrabRelac).

No **grupo (1)**, não há uma especificação sobre o planejamento da avaliação no ciclo de vida de desenvolvimento na maioria das abordagens analisadas. Algumas delas

não incluem a avaliação externa no seu planejamento; ou não especificam o *feedback* durante o jogo; ou estão limitadas ao jogo criado usado para avaliar a abordagem. Por exemplo, Petri (2018) apresenta um modelo que contém o processo de planejamento e condução para avaliar a qualidade de jogos educacionais (em termos de satisfação e usabilidade). No entanto, ele não especifica uma avaliação interna, como registro de interação e coleta de dados de desempenho. No **grupo (2)**, os modelos canvas não especificam a avaliação. Apenas Star *et al.* (2016) contemplam esse suporte, porém o canvas é específico para jogos pró-sociais. No **grupo (3)**, quase todos os jogos analisados são desenvolvidos sem considerar o planejamento da avaliação de desempenho dos alunos. A avaliação da aprendizagem geralmente é realizada por meio de pré e pós-testes, sem avaliações durante o jogo (avaliação formativa e coleta de dados das ações dos jogadores durante o jogo).

### 3. Metodologia de Pesquisa

Este trabalho é caracterizado: quanto à sua natureza como uma pesquisa **aplicada**, por propor o desenvolvimento de um modelo conceitual; e, quanto aos objetivos, como uma pesquisa **descritiva**, por prever a identificação e registro de características de processo de desenvolvimento do modelo proposto, e **exploratória**, por buscar compreender melhor o problema a partir da análise do estudo da arte (Gil, 1991). A pesquisa tem abordagem quali-quantitativa, pois utiliza multimétodo, e foi delineada em etapas, conforme descritas a seguir.

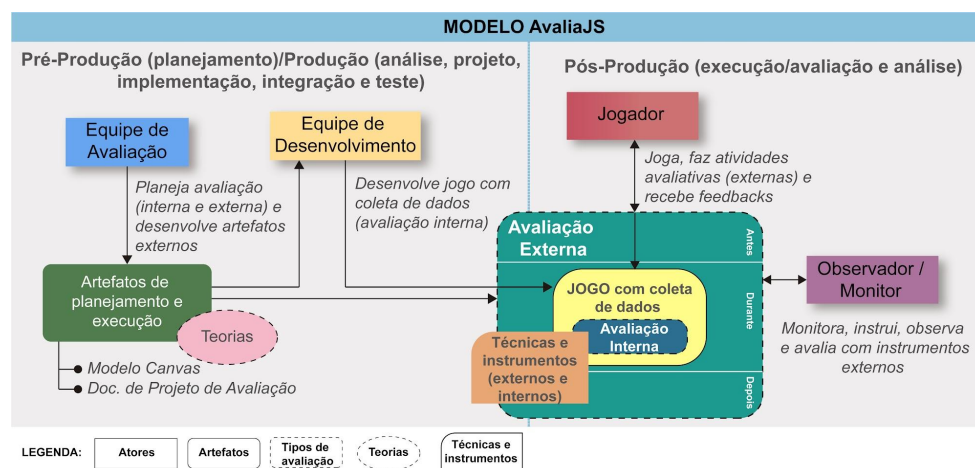
A etapa **Levantamento do estado da arte** objetivou compreender os conceitos mais relevantes, sintetizados por meio de uma pesquisa bibliográfica e realizar o levantamento do estado da arte, por meio de uma revisão *ad-hoc* da literatura para compreensão de trabalhos relacionados. Além disso, uma Revisão Sistemática da Literatura (Oliveira *et al.*, 2018) foi realizada para analisar *frameworks* usados no desenvolvimento de jogos educacionais. Na etapa **Desenvolvimento do modelo**, o AvaliaJS foi desenvolvido a partir de uma visão holística sobre as abordagens de avaliação e as referências de modelos canvas analisadas na etapa anterior. O canvas e o documento de projeto de avaliação foram criados baseados nos métodos de planejamento de projeto (Rossato, 1996; Falchikov, 2005), e estruturados de acordo com as diretrizes de criação e estruturação de canvas (Osterwalder, 2004).

Em relação à avaliação, a etapa **Aplicação e avaliação do modelo** objetivou analisar três jogos prontos para avaliar e exemplificar o uso do modelo conceitual como prova de conceito, no que diz respeito à cobertura de conteúdos para preenchimento do modelo canvas. Os jogos selecionados foram escolhidos por atenderem objetivos definidos (ensino-aprendizagem, treinamento, avaliação, motivação e engajamento), por estarem disponíveis (jogo/código-fonte e documentações) e porque os autores fizeram parte do processo de desenvolvimento dos mesmos. A análise da aplicação do modelo foi discutida a partir das perspectivas: produto, processos, pessoas e projeto (propostas por Aslan e Balci, 2015). Na etapa **Avaliação do modelo por especialistas**, seis especialistas (Beecham *et al.*, 2005) avaliaram o AvaliaJS e seus artefatos. O planejamento dessa avaliação foi estruturado com o método *Goal/Question/Metric* (GQM) (Basili *et al.*, 1994). Foi usado um questionário para coletar os dados, cujas métricas e questões foram baseadas nos trabalhos de Savi *et al.* (2010), Rocha (2014) e Petri (2018). O questionário foi criado por meio das etapas do processo de construção de medidas de avaliação em Treinamento, Desenvolvimento e Educação (Mourão e

Meneses, 2012), e avaliado por análise semântica (Coelho *et al.*, 2020; Zerbini *et al.*, 2012). Além disso, a análise de consistência interna das respostas do questionário (do painel de especialistas) foi realizada pelo cálculo do coeficiente alfa de Cronbach (1951). A partir dos resultados dessas avaliações, o modelo foi analisado e avaliado, sendo discutidas as suas limitações e ameaças à validade (apresentadas na **Seção 5**).

#### 4. AvaliaJS: modelo e artefatos

O modelo conceitual proposto foi denominado AvaliaJS (*Avalia* = *Avaliação* + *JS* = *Jogo Sério*). Ele foi projetado a partir de uma visão holística da literatura sobre as abordagens de avaliação no contexto de JSs. O AvaliaJS visa que os JSs sejam projetados considerando o planejamento e execução da avaliação do desempenho do aluno e o *feedback* imediato e constante. Para isso, em relação aos profissionais envolvidos (conforme ilustrado na **Figura 1**), de modo geral, a **equipe de avaliação** (analista, especialistas no domínio, em avaliação e em jogos) é responsável por planejar a avaliação (externa e interna) usando as teorias adequadas, conforme o foco do JS. Por exemplo, teorias como a taxonomia de bloom, níveis de avaliação de Kirkpatrick, teoria dos erros humanos, dimensões de *feedback*, modelo ARCS e aprendizagem afetiva (Eseryel *et al.*, 2011; Ifenthaler *et al.*, 2012; Kirkpatrick e Kirkpatrick, 2016; Rocha, 2014; Oliveira e Rocha, 2020). Além disso, a equipe de avaliação desenvolve artefatos para avaliação externa, como questionários, entrevistas e *debriefing*. A **equipe de desenvolvimento** (*designer* e programador) é responsável por desenvolver o jogo com arquitetura, estruturas e algoritmos adequados à coleta de dados (avaliação interna). Os artefatos de avaliação externa são projetados para serem usados pelos **jogadores** antes, durante ou depois da aplicação do jogo, e um **observador/monitor** (pesquisador, professor ou instrutor) monitora, instrui, observa e avalia com esses artefatos externos.



**Figura 1: AvaliaJS: modelo conceitual de planejamento do design e execução da avaliação do desempenho de alunos em JSs.**

As **Figuras 1** e **2** apresentam os diagramas que esboçam o AvaliaJS de modo geral e específico, com um recorte da etapa de execução/avaliação, em que são apresentados técnicas e instrumentos de avaliação e coleta de dados.

Em relação ao suporte para criação dos produtos, o AvaliaJS possui dois artefatos, que visam a descrição das teorias, técnicas e instrumentos, tanto externos e internos, que serão usados para a execução da avaliação, são eles: (1) **Modelo canvas**

(planejamento em alto nível) e (2) **Documento de projeto da avaliação** (descrição em baixo nível). O modelo canvas foi desenvolvido baseando-se em questões de planejamento de uma avaliação, adaptadas de Rossato (1996) e Falchikov (2050). As questões são: (i) *O que* será avaliado durante o jogo?; (ii) *Por que* a avaliação deve ser realizada?; (iii) *Como* deverá ser realizada a avaliação?; (iv) *Quando* cada uma das avaliações deverá ser executada?; (v) *Onde* cada avaliação será executada?; e (vi) *Quem* está envolvido na execução das avaliações?. Após a definição das questões, o modelo canvas foi estruturado baseando-se nas características de construção de um *Business Model Canvas* (BMC) (Osterwalder, 2004), como organização das seções por influência, agrupamento por relação e significado atômico das seções. O modelo canvas permite uma rápida identificação dos elementos e atividades necessárias para o planejamento do *design* e execução da avaliação do desempenho dos alunos em JSs (disponível em: [bit.ly/CanvasAvaliaJS](http://bit.ly/CanvasAvaliaJS)). No documento de projeto da avaliação, o que foi planejado no modelo canvas deve ser detalhado (disponível em: [bit.ly/DocProjetoAvaliaJS](http://bit.ly/DocProjetoAvaliaJS)).

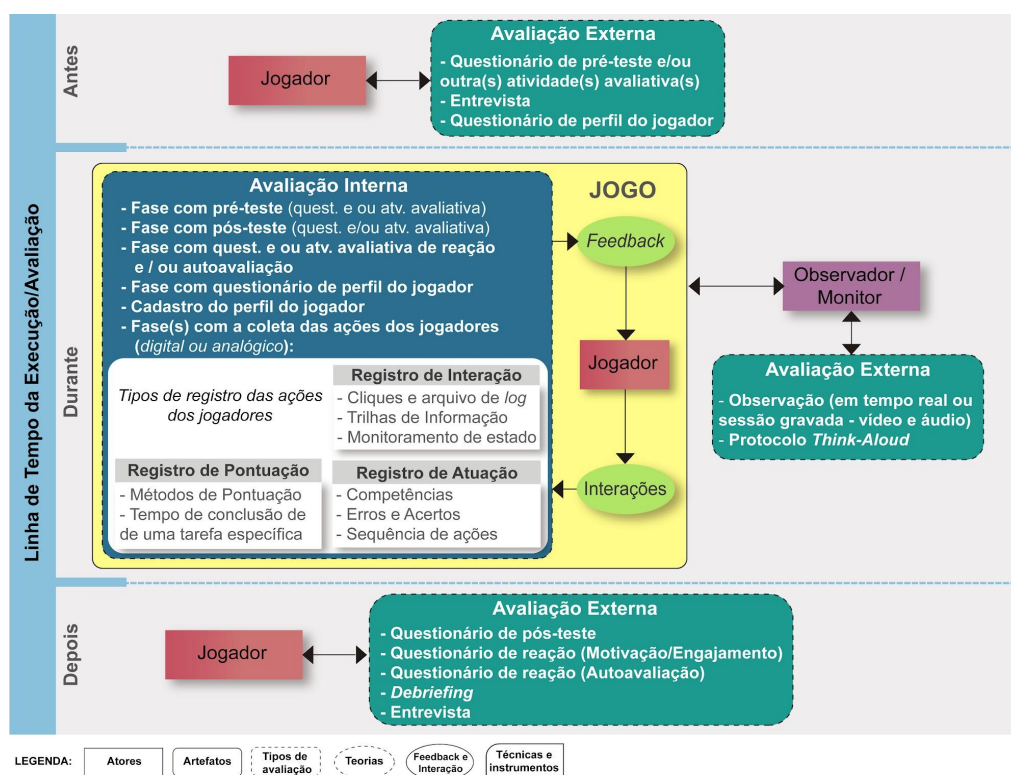


Figura 2: Visão detalhada da etapa de execução/avaliação do modelo AvaliaJS

## 5. Avaliações e Resultados

O AvaliaJS foi avaliado por meio de duas técnicas (prova de conceito e painel de especialistas), cujo objetivo foi definido de acordo com o método GQM (Basili *et al.*, 1994) e os principais resultados são destacados a seguir.

(1) **Prova de conceito:** Analisar os jogos GLPSobControle (Rocha, 2014), Expedição Antártica (Oliveira *et al.* 2019) e Guerra em Alto Mar (Oliveira e Rocha, 2019) com o propósito de avaliar e exemplificar o uso do modelo conceitual como prova de conceito, em relação à cobertura de conteúdos para o preenchimento do modelo, do ponto de vista dos pesquisadores (autores), no contexto de desenvolvimento

e avaliação JSs. O preenchimento do modelo canvas e do documento do projeto foi elaborado pelos pesquisadores (autores que também participaram das produções desses jogos), após a definição dos jogos a serem analisados. Para realizar a avaliação, os pesquisadores responderam à questão: *o modelo inclui e cobre o conteúdo necessário para o planejamento do design e execução da avaliação do jogador no jogo sério (conforme realizado em cada jogo)?* Por se tratar de uma avaliação qualitativa, as métricas foram definidas para identificar os conteúdos para preencher o modelo: focos, propósitos, teorias, técnicas, instrumentos e infraestruturas. Como resultado, foi possível verificar que o modelo cobre os principais conceitos necessários para planejar a avaliação em JSs, sobretudo para os jogos que focam no ensino-aprendizagem, treinamento, motivação e engajamento. Os modelos canvas e os documentos de projeto da avaliação dos jogos exemplificados podem ser consultados em, respectivamente: <https://bit.ly/GLPCanvas>; <https://bit.ly/GuerraCanvas>; <https://bit.ly/AntarticaCanvas>.

(2) **Painel de especialistas:** Analisar o AvaliaJS (diagramas, canvas e documento de projeto) com o propósito de avaliar a qualidade, em relação a sua corretude, consistência, clareza, inequívoco, completude, autenticidade, flexibilidade e usabilidade, do ponto de vista de especialistas em JSs e em avaliação, no contexto de avaliação do desempenho de alunos em jogos. Um protocolo de avaliação (com relatório técnico e questionário) foi enviado por *e-mail* para seis especialistas, disponível em: [bit.ly/ProtocoloAvaliaJS](https://bit.ly/ProtocoloAvaliaJS). Eles foram selecionados por conveniência, por estarem envolvidos no desenvolvimento dos JSs usados para exemplificação do modelo e por terem conhecimento da avaliação do desempenho humano em diferentes perspectivas e áreas. As questões do GQM foram definidas de acordo com os oito critérios de qualidade (Q1 a Q8). Em relação a corretude (Q1), consistência (Q2), clareza (Q3), inequívoco (Q4) e completude (Q5) do modelo AvaliaJS, os especialistas responderam positivamente "sim" aos critérios (cinco questões com escala dicotômica - sim ou não). Em relação à autenticidade (Q6), flexibilidade (Q7) e usabilidade (Q8), os resultados da análise das respostas dos especialistas, para sete assertivas com escala tipo *Likert* (discordo totalmente (1) a concordo totalmente (5)), confirmam uma boa consistência interna, conforme alfa de Cronbach  $\alpha = 0,87$ . O cálculo do alfa de Cronbach está disponível em: [bit.ly/CalcAlfa](https://bit.ly/CalcAlfa). Como resultado, há indícios de que o AvaliaJS é correto, autêntico, consistente, claro, inequívoco, completo e flexível, mas precisa melhorar sua usabilidade.

### 5.1. Limitações e Ameaça à validade

As limitações e ameaças à validade relacionadas a este estudo são apresentadas a seguir, classificadas em quatro categorias de acordo com Wohlin *et al.* (2000).

Na categoria **Validade interna**, O termo “Jogo Sério” é abrangente, pois envolve diferentes finalidades e domínios, então, é possível que isso pode ameaçar o escopo do modelo. Para minimizar possíveis problemas, limitou-se o escopo do modelo a JSs com foco em educação e treinamento/simulação. Além disso, a aderência do modelo às diferentes teorias, técnicas e instrumentos para a avaliação nos JSs foi verificada apenas nos três jogos utilizados como prova de conceito. Em relação à avaliação com especialistas, é possível que eles possam ter sido afetados pelo prazo de avaliação do modelo. Para amenizar isso, eles receberam a documentação do AvaliaJS por *e-mail* e puderam avaliar no momento mais conveniente. Outro problema poderia ser o número de documentos que os especialistas tiveram que avaliar. Para minimizar

este problema, elaboramos um relatório técnico e um protocolo com tarefas bem definidas. Enquanto que na categoria **Validade Externa**, uma ameaça à possibilidade de generalização dos resultados está relacionada ao número e à experiência dos especialistas. Quanto ao tamanho ( $N = 6$ ), esse número pode ser utilizado para avaliar e apoiar o desenvolvimento de modelos ou conceitos (Beecham *et al.*, 2005). Além disso, todos são especialistas envolvidos na produção dos jogos usados, o que facilitou a avaliação da fidelidade dos artefatos. Entretanto, não podemos generalizar essa avaliação para todos os contextos, *e.g.*, para não especialistas.

Na categoria **Validade do Construto**, quanto à interpretação do preenchimento do modelo (como prova de conceito), a possibilidade de não compreensão dos três jogos usados foi mitigada pelo envolvimento dos pesquisadores (autores) na produção desses jogos. Para minimizar a possibilidade de não compreensão do questionário usado no painel pelos especialistas, o questionário foi elaborado considerando os processos de construção de medidas de avaliação e avaliado por meio de análise semântica e sintática com dois especialistas em jogos e tecnologia da informação na educação. Além disso, a confiabilidade do questionário (respostas desse painel) foi analisada por meio do coeficiente de Cronbach, indicando uma boa consistência interna. Por fim, na categoria **Validade de Conclusão**, o modelo não foi avaliado na criação de um novo jogo. Assim, os resultados não podem ser considerados conclusivos, mas sim um indicativo de que o modelo é funcional e abrange os conteúdos que podem ser utilizados para avaliação em JSs no âmbito deste estudo.

## 6. Considerações Finais e Trabalhos Futuros

Os jogos sérios são desenvolvidos com o objetivo principal de avaliar o progresso e os resultados do aprendizado. No entanto, muitos desses jogos não são projetados com avaliação de desempenho e *feedback* eficazes. Esse problema pode ser devido à falta de tempo de desenvolvimento, falta de envolvimento de especialistas do domínio no processo de produção ou foco na transmissão de conteúdo em vez de avaliação. Com o objetivo de tratar esse problema, este trabalho relatou a criação, descrição e avaliação do AvaliaJS, um modelo conceitual que auxilia no planejamento do *design* e execução da avaliação do desempenho de alunos em JSs. A avaliação por prova de conceito e painel de especialistas demonstra que o AvaliaJS atende aos critérios de qualidade: corretude, consistência, clareza, inequívoco, completude, autenticidade, flexibilidade. No entanto, precisa melhorar a sua usabilidade, visto que os especialistas não avaliaram o modelo usando-o para planejar a avaliação em um jogo.

Ressalta-se a originalidade do trabalho por apresentar um modelo canvas com foco na avaliação, uma vez que a maioria dos canvas são para *game design*. Além disso, o modelo tem o foco não apenas na avaliação da reação mas, principalmente, na avaliação da aprendizagem de alunos em JSs. Destaca-se a relevância do uso de um modelo conceitual como o AvaliaJS no desenvolvimento do JS, para planejamento, documentação e efetiva comunicação com os outros profissionais envolvidos e assim inclusão da coleta de dados, avaliação do desempenho humano e *feedback*. Além disso, o modelo AvaliaJS pode impactar na redução de retrabalho e de custos na produção de JSs; e, como impacto social, fornecer uma avaliação efetiva, com *feedback* construtivo durante a execução de tarefas e, assim, contribuir para a aprendizagem dos alunos e, conseqüentemente, na qualidade do ensino por meio de JSs.

Este trabalho contribuiu para as áreas de Engenharia de *Software* e Informática na Educação (mais especificamente o uso de Jogos Sérios na Educação) pela produção de artefatos teóricos (relatório e exemplificações de uso), científicos (publicações de quatro artigos completos, um artigo resumido, um relatório técnico e um artigo completo em processo de revisão) e tecnológico (modelo conceitual e seus diagramas):

(1) *Criação e avaliação de um modelo conceitual (AvaliaJS)*: contribuição tecnológica, principalmente para a área de Engenharia de *Software*, o *AvaliaJS* foi relatado e discutido sua avaliação por um painel de especialistas. Os autores foram convidados a submeterem uma versão estendida para **JIS (Qualis B3)**, na *Special Issue Best of SBGames 2020 - Education and Health* (foi submetido e encontra-se em processo de revisão) (Oliveira *et al.*, 2021b);

(2) *Usos do AvaliaJS (para exemplificação e avaliação do modelo)*: (i) revisão e avaliação de um jogo de tabuleiro (resumo expandido publicado no **SBGames - Qualis B2**) (Oliveira e Rocha, 2019); (ii) desenvolvimento e avaliação de um JS acessível (artigo publicado na revista **RBEE - Qualis A2**) (Oliveira *et al.*, 2021a); e (iii) descrição e discussão do uso de três jogos já prontos (artigo publicado no **SBGames - Qualis B2**) (Oliveira e Rocha, 2020b).

(3) *Criação de instrumentos de avaliação (da reação, aprendizagem e comparação de jogos)*: que foram propostos e usados na avaliação de jogos educacionais (artigo publicado no **SBIE - Qualis B1**) (Oliveira *et al.*, 2019);

(4) *Resumo e análise do estado da arte*: dos principais *frameworks* conceituais, levantados por meio de uma revisão sistemática da literatura, usados no desenvolvimento de jogos/JSs e das lacunas na área de avaliação em JSs (artigo publicado no **SBIE - Qualis B1**) (Oliveira *et al.*, 2018);

(5) *Compilação e organização de referencial teórico*: que fornece uma base para os diferentes profissionais envolvidos, em um relatório técnico sobre o uso do modelo *AvaliaJS* (Oliveira e Rocha, 2020a), visto a necessidade da compreensão da área, para criação de JSs que contemplem a avaliação do desempenho do aluno e *feedback*. No relatório, há um glossário com a descrição dos conteúdos a serem alocados em cada seção no *AvaliaJS* e suas respectivas referências.

Como trabalhos futuros, pretende-se validar o modelo na criação de um novo jogo e o *AvaliaJS* deve ser aprimorado a partir da sua integração em processos bem definidos do ciclo de vida de produção do jogo e especificação das funções, papéis e atores, atividades e artefatos de entrada e saída. Pretende-se criar e validar uma ferramenta computacional para incluir, customizar e analisar a avaliação e geração de relatórios dos dados coletados. Além disso, o modelo deve ser expandido para incluir os custos de avaliação do projeto e execução (como valor dos materiais, instrumentos e artefatos usados); as questões de confiabilidade da avaliação (técnicas de validação da avaliação e tipos de confiabilidade) e a análise dos resultados (técnicas para apresentar os resultados e refletir os resultados educacionais alcançados).

## Agradecimentos

Este trabalho foi realizado com o apoio financeiro da PROEC/UFABC e CAPES (Processo 88887.361026/2019-00).

## Referências

Abt, C. C. (1987). "Serious Games". Lanham. MD: University Press of America.



- Aslan, S.; Balci, O. (2015). GAMED: digital educational game development methodology. In "Simulation", v. 9(4), p. 307-319.
- Basili, V.; Caldiera, G.; Rombach, H. (1994). "Goal Question Metric Paradigm". Encyclopedia of Software Engineering, v. 1, John Wiley & Sons, p. 528-532.
- Beecham, S.; Hall, T.; Britton, C.; Cottee, M.; Rainer, A. (2005). "Using an Expert Panel to Validate a Requirements Process Improvement Model". The J. Syst. Soft. v. 76.
- Bellotti, F.; Kapralos, B.; Lee, K.; Moreno-ger, P.; Berta, R. (2013). "Assessment in and of serious games: an overview". Advances in Hu.-Comp. Inter., v. 2013, p. 1-11.
- Chaudy, Y; Connolly, T. (2019). "Specification and evaluation of an assessment engine for educational games: Integrating learning analytics and providing an assessment authoring tool". In: Entertainment Comput., v. 30, p. 1-16.
- Coelho, J.A.P.M.; Souza, G.H.S.; Albuquerque, J. (2020). Desenvolvimento de questionários e aplicação na pesquisa em Informática na Educação. Jaques, P. A. et al., (Org.) "Metodologia de Pesquisa Científica em Informática na Educação: Abordagem Quantitativa". Porto Alegre, RS: SBC.
- Cronbach, L. J. (1951). "Coefficient alpha and the internal structure of tests". Psychometrika, v. 16(3), p. 297-334.
- Daoudi, I.; *et al.* (2017). "Learners' Assessment and Evaluation in Serious Games: Approaches and Techniques Review". ISCRAM-med, Xanthi, Greece.
- Emmerich, K.; Bockholt, M. (2016). "Serious Games Evaluation: Processes, Models, and Concepts". Entert. Comp. and Serious Games. Springer, Cham, v. 9970, p. 265-283.
- Falchikov, N. (2005). "Improving Assessment Through Student Involvement: Practical Solutions For Aiding Learning In Higher And Further Education". RoutledgeFalmer.
- GIL, A. C. (1991). "Como elaborar projetos de pesquisa". 3 ed. SP: Atlas, p. 1-192.
- Hattie, J.; Timperley, H. (2007). "The power of feedback". Review of Educational Research, v. 77(1), p. 81-112.
- Kirkpatrick, J. D.; Kirkpatrick, W. K. (2016). "Kirkpatrick's Four Levels of Training Evaluation". ATD Press.
- Mourão, L.; Meneses, P. P. M. (2012). Construção de Medidas em TD&E. In "Medidas de Avaliação em Treinamento, Desenvolvimento e Educação: ferramentas para gestão de pessoas". PA, RS: Artmed, p. 50-63.
- Oliveira, R. N. R. *et al.* (2019). "Avaliações em Jogos Educacionais: instrumentos de avaliação da reação, aprendizagem e comparação de jogos". In: SBIE, p. 972-981.
- Oliveira, R. N. R. *et al.* (2021a). "Relato de Experiência do Desenvolvimento e Avaliação da Usabilidade e Acessibilidade de um Protótipo de Jogo Educacional Digital para Pessoas com Deficiência Visual". In: RBEE, v.27(3), p. 847-864.
- Oliveira, R. N. R.; Cardoso, R. P.; Braga, J. C.; Rocha, R. V. (2018). "Frameworks para Desenvolvimento de Jogos Educacionais": uma revisão e comparação de pesquisas recentes". In: SBIE, p. 854-863.
- Oliveira, R. N. R.; Rocha, R. V.; Goya, D. H. (2021b). "Planning the Design and Execution of Student Performance Assessment in Serious Games". In: Journal on

- Interactive Systems (JIS) - Special Issue Best of SBGames 2020 - Education and Health, p. 1-15.
- Oliveira, R.N.R.; Rocha, R.V. (2019). “Guerra em Alto Mar: um Jogo de Tabuleiro com Quiz Personalizável para Engajar e Motivar Estudantes”. In: SBGames, p. 1164-1167.
- Oliveira, R.N.R.; Rocha, R.V. (2020a). “AvaliaJS: Modelo Conceitual de Planejamento da Avaliação do Desempenho de Alunos em Jogos Sérios”. Relatório Técnico. UFABC, 2020a. Disponível em: <https://bit.ly/RelatorioJS>. Acesso em: maio. 2020.
- Oliveira, R.N.R.; Rocha, R.V. (2020b). “Modelo Conceitual de Planejamento da Avaliação do Desempenho de Alunos em Jogos Sérios”. In: SBGames, p. 682-691.
- Oliveira, W.; Joaquim, S. and Isotani, S. (2020). Avaliação de Jogos Educacionais: desafios, oportunidades e direcionamento de pesquisa. In *XIX SBGames*, p. 775-778.
- Osterwalder, A. (2004). “The Business Model Ontology: A Proposition in a Design Science Approach”. Thesis (Bus. Infor. Syst.) - Univ. of Lausanne, Switzerland.
- Petri, G. (2018). A method for the evaluation of the quality of games for computing education. Dissertation (Computer Science) - Federal University of Santa Catarina. Florianópolis, SC, p. 1-335.
- 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) – UFSCar.
- Rocha, R.V.; Bittencourt, I.I.; Isotani, S. (2015). “Análise, Projeto, Desenvolvimento e Avaliação de JSs e Afins: uma revisão de desafios e oportunidades”. In SBIE, p. 692-701.
- Rossato, I. F. (1996). “Uma metodologia para a análise e solução de problema”. Dissertação (Mestrado em Engenharia de Produção e Sistema) - UFSC.
- Salas, E. *et al.* (2009). Performance measurement in simulation-based training: a review and best practices. In *Simulation & Gaming*, v. 40, p. 328–376.
- Savi, R.; Gresse Von Wangenheim, C.; Ulbricht, V. R.; Vanzin, T. (2010). “Proposta de um modelo para avaliação de jogos educacionais”. In RENOPE, v. 8(3), p. 1-10.
- Star, K.; Vuillier, L.; Deterding, S. (2016). D2. 6 Prosocial Game design methodology. Gamification of Prosocial Lear. for Incr.Youth Inclusion and Acad. Achi., p. 1-60.
- Trybus, J. (2010). “Game-Based Learning: What it is, Why it Works, and Where it's Going”. NMI White Paper. New Media Institute, New York.
- Wohlin, C.; et al. (2000). “Experimentation in Software Engineering-An Introduction”. (1st ed.). Kluwer Academic Publishers, Norwell, MA, USA.
- Zerbini, T.; Borges-ferreira, M. F.; Abbad, G. S. (2012). Medidas de Reação a Cursos a Distância. In “Medidas de Avaliação em Treinamento, Desenvolvimento e Educação: ferramentas para gestão de pessoas”. PA, RS: Artmed, p. 91-107.