

Avaliação de conhecimento em jogos sérios: Uma revisão sistemática de literatura

Maiara Miranda¹, Michelle Nery Nascimento¹, Gerda de Oliveira¹,
João Pereira¹, Lucila Ishitani¹

¹Programa de Pós-Graduação em Informática
Instituto de Ciências Exatas e Informática (PUC Minas) – Belo Horizonte – MG – Brasil

maiara.fernanda@sga.pucminas.br, michellenerly@pucminas.br
{gerda.graciela, joao.pereira.1129311}@sga.pucminas.br, lucila@pucminas.br

Abstract. *Serious games are used to teach and provide learning experiences, with applications in areas such as education, healthcare, government and corporations. However, to verify the learning obtained through these games, it is necessary to carry out evaluations. This article brings a systematic review of the literature to investigate the methods used to evaluate the gain of knowledge through educational games, a total of 116 articles were selected and were found evaluation methods such as pre-test, post-test, questionnaire, observation and in-game monitoring. The synthesis of these methods provides guidance for serious game developers, however, in future research, it may be interesting to further explore in-game evaluation strategies.*

Resumo. *Jogos sérios são utilizados para transmitir mensagens, ensinar e fornecer experiências de aprendizado, com diversas aplicações em áreas como educação, saúde, governo e corporações. Porém, para verificar o aprendizado obtido por meio desses jogos é necessário realizar avaliações, este artigo traz uma Revisão Sistemática de Literatura para investigar os métodos usados para avaliação do ganho de conhecimento por meio de jogos educativos. Foram selecionados 116 artigos no total e encontradas práticas de avaliação que incluíam pré-teste, pós-teste, questionário, observação e monitoramento. A síntese desses métodos fornece orientações para a comunidade desenvolvedora de jogos sérios, porém em futuras pesquisas pode ser interessante aprofundar a estratégias de avaliação in-game.*

1. Introdução

Jogos sérios são usados como ferramentas para o aprendizado e treinamento em uma variedade de contextos, no entanto, o uso desses jogos não garante resultados positivos de aprendizado. É por isso que a realização de avaliações adequadas se torna essencial, pois elas permitem verificar se os jogadores estão de fato adquirindo conhecimento e/ou desenvolvendo habilidades relevantes.

A avaliação de aprendizado em jogos sérios fornece dados concretos sobre a eficácia da abordagem educacional adotada, possibilitando melhorias contínuas e aprimoramento da experiência de aprendizado dos jogadores. Em outras palavras, a avaliação de aprendizado e do impacto dos jogos sérios fornece informações valiosas para os desenvolvedores, educadores e profissionais envolvidos, para que seja possível identificar

áreas de aprimoramento, como a necessidade de adequação da forma de apresentação de informações ou de adaptação do conteúdo às necessidades e características do público-alvo.

Neste trabalho foi realizada uma Revisão Sistemática de Literatura para investigar quais os métodos de avaliação de ganho de conhecimento por meio de jogos educacionais têm sido utilizados. Na literatura, os métodos com mais ocorrências foram: pré-teste e pós-teste, questionário, monitoramento e avaliação *in-game* e combinações desses métodos.

As seções deste trabalho detalham o processo adotado e seus resultados. A Seção 2 apresenta a conceituação de jogos sérios identificada na revisão bibliográfica, seguida da Seção 3 que detalha a metodologia utilizada para seleção e análise dos artigos selecionados sobre avaliação de aprendizagem. A Seção 4 mostra os resultados da Revisão Sistemática de Literatura. Por último, as seções 5 e 6 apresentam, respectivamente a discussão e as conclusões deste trabalho.

2. Revisão bibliográfica

2.1. *Serious games*

De acordo com Michael e Chen (2005), jogos sérios são jogos que usam a narrativa e o mundo dos jogos para transmitir uma mensagem, ensinar ou fornecer uma experiência. Atualmente, os jogos sérios têm sido cada vez mais empregados em diversas áreas, como engenharia, turismo, medicina, construção civil, educação, entre outras. Esse crescimento se deve, em parte, à redução dos riscos de exposição a situações perigosas do mundo real e aos custos mais baixos de treinamento prático.

Ainda de acordo com Michael e Chen (2005), todos os jogos são uma forma de expressão. Como os jogos expressam ideias, informações e crenças, isso significa que os jogos ensinam.

A principal característica de jogos sérios é a sua finalidade educacional, não tendo como objetivo primário ser utilizado para a diversão, o que não significa que jogos sérios não possam ou devam ser divertidos. O uso de jogos como recurso na educação possui algumas vantagens, pois eles são produtos interativos e motivadores, despertando o interesse e a motivação dos estudantes por meio de desafios, curiosidade e interação. Além disso, facilitam o aprendizado ao utilizar recursos gráficos que auxiliam na compreensão dos temas abordados, aumentando a interação e o interesse [Savi and Ulbricht 2008].

Os jogos sérios também promovem o desenvolvimento de habilidades cognitivas, pois exigem estratégias, tomada de decisões e compreensão dos elementos e interações do jogo. Eles incentivam a aprendizagem por descoberta, explorando ambientes virtuais livres de perigos reais. Eles também contribuem para o desenvolvimento da coordenação motora e habilidades espaciais por meio da interação com controles e exploração do ambiente virtual.

Savi e Ulbricht (2008) descrevem em seu trabalho que os jogos sérios possuem algumas características que os tornam excelentes ferramentas de auxílio no processo ensino-aprendizagem, sendo elas:

- Efeito motivador por provocar o interesse e motivação nos estudantes com desafios, curiosidade e interação, ao mesmo tempo que incentivam o aprendizado;

- Facilitador do aprendizado por utilizar recursos gráficos que facilitam o entendimento de certos assuntos;
- Desenvolvimento de habilidades cognitivas;
- Aprendizado por descoberta, pois incentivam a capacidade exploratória e colaborativa dos jogadores;
- Experiência de novas identidades por meio dos avatares dos jogadores e imersão em novos mundos;
- Socialização pela possibilidade de aproximação dos estudantes de forma competitiva ou cooperativa, dentro do mundo virtual;
- Coordenação motora e habilidades espaciais, pela exploração do ambiente virtual.

Jogos sérios não possuem uma classificação única e definida na literatura, mas Michael e Chen (2005) os classificam como:

Jogos militares - foram um dos primeiros tipos de jogos sérios a surgir na década de 1970. São jogos utilizados para treinamento de tropas e simulações de conflitos reais.

Jogos governamentais - são jogos que não se restringem a questões militares e envolvem diversos níveis do governo, nacional estadual e municipal.

Jogos educativos - são concebidos com o propósito de promover a aprendizagem, buscando conciliar de forma harmoniosa o conteúdo abordado com a experiência de jogo e a capacidade do jogador de assimilar e aplicar o conhecimento adquirido no mundo real.

Jogos corporativos - os jogos corporativos deslocam o foco do setor público para o setor privado, explorando o uso de jogos sérios no ambiente corporativo. Empresas de todos os tamanhos estão investindo em jogos sérios como ferramentas tanto para treinamento quanto para planejamento estratégico.

Jogos para a área da saúde - os jogos de saúde são uma forma de jogos sérios que encontram aplicação na área da saúde, abrangendo tratamento, recuperação e reabilitação. Além disso, o uso de jogos sérios na saúde também se estende ao campo da saúde mental.

Jogos políticos, religiosos e artísticos - jogos políticos têm como foco narrativas sobre tensões políticas não resolvidas e, normalmente, tendem a estimular controvérsia. Os jogos religiosos têm por objetivo ensinar ou divulgar conteúdo religioso e jogos artísticos objetivam expressar ideias artísticas ou arte no mundo do jogo.

Para Santos e Pereira (2020), a aprendizagem lúdica tem-se tornado uma ferramenta poderosa na área da aprendizagem pois ela permite a construção do conhecimento através da satisfação e da diversão do indivíduo, tornando-se um elemento facilitador do aprendizado.

Por fim, de acordo com De Freitas e Liarokapis (2011), o potencial dos jogos sérios irá transformar o paradigma e a maneira como a educação e o treinamento são conduzidos até os dias atuais.

2.2. Trabalhos relacionados

Enquanto este estudo foca nos métodos para avaliar o aprendizado conquistado através de jogos sérios, Krath et al. (2021) realizaram uma revisão sistemática de literatura focada na teoria por trás dos jogos sérios e da gamificação. A pesquisa foi feita nas bases Web of Science Core Collection, EBSCO Host, Wiley Online, EmeraldInsight, JSTOR, SagePub

e Taylor e Francis, além da IEEE Explore e Science Direct. Diferentemente deste trabalho, os autores tiveram o objetivo de responder à pergunta “Qual a fundamentação teórica usada em pesquisas sobre gamificação, jogos sérios e aprendizado através de jogos”.

Por sua vez, Ullah et al. (2022) conduziram uma revisão de literatura focada apenas na educação científica e de forma mais ampla, em busca das tendências e padrões de pesquisa sobre o assunto. Apesar de buscarem entender de que maneira a intervenção feita por jogos sérios tem sido avaliada, esse não é o foco do artigo, que traz discussões sobre o uso de jogos sérios de modo geral e seus aspectos positivos e negativos para a educação científica.

Baseados nas questões “videogames podem melhorar habilidades cognitivas e de decisões?” e “quais gêneros de videogames afetam habilidades cognitivas?”, Reynaldo et al. (2021) estudaram se de fato jogos sérios trazem uma contribuição para o usuário, mostrando superficialmente se, e qual, método foi usado para avaliar o conhecimento ganho a partir do uso de jogos.

Petri e von Wangenheim (2016) realizaram uma revisão sistemática de literatura em como avaliar jogos educacionais buscando artigos de até 2015, nas bases Wiley Online, Springer Link, ACM, IEEE e Science Direct. Diferentemente deste trabalho, as perguntas de pesquisa envolviam como o próprio método de avaliação foi avaliado e desenvolvido. Tendo em vista como o universo de jogos educacionais vêm crescendo, há outros sendo atualizados nesses 7 anos de intervalo entre este estudo e o artigo de Petri e von Wangenheim.

3. Metodologia

Neste estudo foi realizada uma Revisão Sistemática de Literatura (RSL) [Kitchenham et al. 2009] com o objetivo de responder às seguintes questões de pesquisa:

- Q1** Como foi avaliado o ganho de conhecimento por meio de jogos educativos?
 - Q1.1** Quais instrumentos foram adotados na avaliação?
 - Q1.2** Quais métodos de avaliação de aprendizado envolveram a participação do público-alvo?
 - Q1.3** Quais métodos de avaliação de aprendizado NÃO envolveram a participação direta do público-alvo?

Os artigos foram pesquisados entre 20 e 24 de setembro de 2022 nas seguintes bases: ACM Digital Library, IEEE EXplore, Science Direct (Elsevier) e Google Acadêmico (*Scholar Google*).

A *string* de busca investigada inicialmente, com termos em inglês e português, foi: “jogo sério” AND “avaliação da aprendizagem”. Devido ao número limitado de artigos obtidos desta forma, foi adicionado também o termo “jogo educativo”. Dessa forma, a *string* ficou definida como (“jogo sério” OR “jogo educativo”) AND avaliação, ou, em inglês, (“serious game” OR “educational game”) AND evaluation. A pesquisa foi restrita ao resumo dos artigos.

Como critério de inclusão foram selecionados apenas os artigos em português ou inglês, com acesso livre ou disponível pelo portal de periódicos da Capes, que efetivamente abordassem uma avaliação de aprendizagem. Por sua vez, os critérios de

exclusão foram: artigos que abordavam apenas avaliação da experiência do usuário, interação humano-computador, aprendizado de máquina ou avaliavam apenas o método de avaliação e não a aquisição de conhecimento em si. Esse processo de seleção de artigos foi executado em duas etapas: primeiro, aplicando os critérios de inclusão e exclusão apenas no título e no resumo dos artigos, mantendo também os artigos cujos título e resumos geravam dúvidas quanto ao foco do trabalho, e, segundo, com uma nova triagem por meio da leitura completa do artigo. Nesta segunda etapa também foram avaliados os critérios de qualidade: se a metodologia de avaliação de aprendizado foi apresentada com detalhes e com clareza.

4. Resultados

A busca inicial resultou em 984 artigos disponíveis para leitura. Em seguida, foi realizada a primeira etapa de seleção de artigos. Nessa primeira etapa foram considerados apenas o título e o resumo dos trabalhos, resultando em 422 artigos. Após, foi feita a leitura completa desses artigos, resultando em um total de 116 artigos.

Tabela 1. Resultados da seleção de artigos

Base	Inicial	Etapa 1	Etapa 2
ACM	136	47	17
Google Acadêmico	99	82	24
IEEE	190	90	39
Science Direct	559	203	36
Total	984	422	116

4.1. RQ1 Como foi avaliado o ganho de conhecimento por meio de jogos educativos?

Buscou-se identificar, nos artigos selecionados, quais métodos ou ferramentas foram utilizados para avaliar a aprendizagem dos jogadores. A Tabela 2 apresenta a quantidade de vezes que cada método de avaliação foi utilizado. Conforme mostrado na tabela, diversos métodos de avaliação foram encontrados nos artigos analisados. Alguns, como a entrevista [Lee et al. 2014] e atividades práticas [Guo and Goh 2016], foram pouco utilizados. Também houve a estratégia de combinar dois ou mais métodos de avaliação, como no trabalho de Robertson et al. (2018). Importante observar que, devido a essas combinações, um mesmo método pode ser citado em mais de uma linha. Por exemplo, o pós-teste é citado em três linhas diferentes, porque ele foi utilizado 4 vezes como recurso único, 46 vezes em conjunto com pré-teste e 7 vezes em conjunto com questionário e pré-teste.

Também foi observado, em alguns artigos como de Chittaro (2016), Desmet et al. (2018) e Hooshyar et al. (2021), que houve uma divisão entre os participantes do estudo em dois grupos: o grupo de estudo (ou experimental) e o grupo controle. O grupo controle é, geralmente, formado pela mesma quantidade e apresenta as mesmas características dos indivíduos presentes no grupo de estudo. Essa divisão faz com que seja possível estudar um devido aspecto de maneira isolada, através da observação dos contrastes entre o grupo que recebeu o tratamento e o grupo que não o recebeu. Ao final do estudo, no momento de comparar os resultados dos dois grupos, é possível constatar que as diferenças se dão por conta do experimento ministrado e não por variáveis externas.

Tabela 2. Método de avaliação e número de ocorrências

Método de avaliação	Quantidade
Pré e pós-teste	46
Questionário	40
Questionário e Pré e pós-teste	7
Monitoramento e avaliação in-game	5
Pós-teste	4
Outros	14

Os métodos mais utilizados são detalhados nas subseções que se seguem.

4.1.1. Pré e pós testes

O método mais utilizado entre os artigos estudados é o pré-teste e pós-teste, que pode ser visto, por exemplo, em Ismail et al. (2016), Ozgen et al. (2020) e Yasin et al. (2019). Este método consiste na aplicação de um exame de conhecimentos anterior ao uso do jogo e, depois, na reaplicação desse mesmo exame (ou de um exame similar) após o uso do jogo educativo. Desta forma, é possível comparar os resultados de ambos os exames e medir a melhora das pontuações. Esta melhora, quando ocorre, é creditada ao processo de aprendizado a partir do uso do jogo, uma vez que este ocorre no intervalo entre a aplicação dos dois exames.

Importante destacar que o momento de aplicação do pós-teste pode variar, podendo ocorrer logo após o uso do jogo ou algum tempo depois. A ideia por trás de realizar o pós-teste dias depois do uso do jogo é para verificar se realmente houve aprendizado, ou seja, se o jogador reteve o conhecimento.

O uso de pré-teste e pós-teste, apesar de muito utilizado, também sofre algumas críticas quanto à sua validade: afinal, quando um mesmo teste é aplicado novamente com os mesmos indivíduos, é esperado que haja uma melhora nos resultados desse teste por conta da repetição.

Artigos como o de Arawjo et al. (2017), von Wangenheim et al. (2012) e Winter et al. (2020) utilizaram, como forma de avaliação, a aplicação de apenas um pós-teste. Entretanto, a ausência de um pré-teste pode fazer falta, pois não é possível afirmar se o indivíduo adquiriu conhecimento com o jogo ou se já possuía aquele conhecimento específico antes do uso do jogo.

4.1.2. Questionário

O segundo método de avaliação mais utilizado é o questionário, que consiste na aplicação de um conjunto de perguntas relativas ao conteúdo do jogo e/ou às impressões do usuário em relação ao seu próprio processo de aprendizado, como observado em Chrysafiadi et al. (2019), Sinclear et al. (2022) e Chang et al. (2017).

Alguns artigos utilizaram questionários prontos, como é o caso do artigo de Schoffel et al. (2018), que utilizou o modelo de avaliação MEEGA+. Já outros artigos desen-

volveram seu próprio questionário de avaliação baseando-se em modelos já conhecidos, como é o caso de Kechai e Pierrot (2015), que utilizaram o modelo TAM (*Technology Acceptance Model*), e de Wang et al. (2018), que utilizaram o modelo ARCS (*Attention, Relevance, Confidence and Satisfaction*).

Entretanto, utilizar apenas um questionário como método de avaliação pode não ser muito eficaz na avaliação de aprendizado, pois depende totalmente das respostas dadas pelo usuário, que podem ter sido dadas sem compromisso, ou podem indicar percepção incorreta de aprendizado.

4.1.3. Monitoramento e avaliação *in-game*

Um outro tipo de avaliação de aprendizado é o monitoramento das atividades e dos resultados do usuário durante a execução do experimento, ou seja, avaliação '*in game*' como ocorre nos trabalhos de Liu e Liu (2020), Harpstead e Alevén (2015) e Qin et al. (2010). Neste método, o conhecimento ganho pelo jogador é avaliado a partir da observação de um avaliador ou da coleta de dados no próprio jogo. No trabalho de Harpstead e Alevén (2015), o jogo automaticamente salva as informações de cada erro, acerto e tempo de resposta para analisar os resultados a partir do método "*Empirical Learning Curve Analysis*". Por sua vez, Qin et al. (2010) analisam essas mesmas informações a partir de testes estatísticos de comparação de amostras Kruskal-Wallis e Wilcoxon-Mann-Whitney.

Esta forma de avaliação ainda não é amplamente usada em jogos sérios, provavelmente pela maior dificuldade de implementação, sendo necessário adicionar esses protocolos na programação do jogo ou designar uma pessoa para acompanhar cada participante durante estudo. Porém, ela é particularmente interessante, pois é capaz de coletar dados que só são possíveis de serem obtidos por meio do acompanhamento do jogador durante a sessão de jogo. Esta maneira de se avaliar se mostrou mais comum em jogos sérios direcionados para a área da saúde, especialmente naqueles em que o objetivo do jogo era o treinamento para alguma cirurgia ou procedimento médico, como em Qin et al. (2010), pois nesses casos se torna importante analisar todo o desempenho do jogador, e não apenas o resultado final.

4.2. RQ1.1 - Avaliação com a participação do público-alvo

Quanto ao público-alvo de cada artigo observou-se que quase em sua totalidade as avaliações tiveram a participação do público-alvo para o qual o jogo foi desenvolvido, com apenas uma exceção. Vale ressaltar que, em 7 artigos, para além da participação do aprendiz houve a participação de outros atores envolvidos, como professores e especialistas ([Duin et al. 2013], [Chrysafiadi et al. 2022], [Rodríguez-Cerezo et al. 2014], [Maertens et al. 2014]). O método de avaliação nesses casos variou entre questionários, pré e pós-teste e entrevistas. A Tabela 3 demonstra, de acordo com o tipo de avaliação, se houve ou não a participação direta do público-alvo.

4.3. RQ1.2 - Avaliação sem a participação direta do público-alvo

Em apenas um artigo selecionado não houve participação direta do público-alvo. Trata-se de um artigo que avaliou um jogo sério desenvolvido para crianças com dificuldades de aprendizagem (crianças com dislexia). Para avaliação foram selecionadas onze crianças

Tabela 3. Participação do público alvo em diferentes tipos de avaliação

Tipo de avaliação	Não	Sim	Sim e Não	Total
Atividade Prática		1		1
Avaliação in-game		4		4
Entrevista		1		1
Entrevista e análise de log		1		1
Fórum de discussões e log de usuário		1		1
Monitoramento de atividades		1		1
Pós-testes		4		4
Pré-teste e Pós-teste		46		46
Pré-teste, pós-teste e entrevista		1	1	2
Pré-teste, pós-teste e observação		1		1
Questionário e avaliação de especialista		1		1
Questionário e entrevista		3		3
Questionário	1	35	4	40
Questionário e Discussões		1		1
Questionário e pré-teste e pós-teste		6	1	7
Questionário, observação e entrevista			1	1
Questionário, observação e pós-teste		1		1
Total Geral	1	108	7	116

com dislexia, com idade entre 8 a 12 anos e especialistas na área de educação especial. O método de avaliação utilizado foi um questionário aplicado a esses especialistas da área de fonoaudiologia. O *feedback* obtido das fonoaudiólogas indicam que o jogo pode favorecer o processo de aprendizagem e ajudar a aprender de forma eficaz [Ouherrou et al. 2018].

4.4. Áreas de aplicação dos jogos abordados nos artigos selecionados

A aplicabilidade dos jogos sérios é ampla, e eles podem ser usados para ensinar temas diversos, desde matemática e ciências a história e línguas. Eles também podem ser usados para ensinar competências sociais, como trabalho em equipe e resolução de conflitos. Além disso, eles podem ser usados para desenvolver habilidades de liderança, comunicação e tomada de decisão.

As áreas de aplicação dos jogos encontrados neste estudo incluem apoio à educação formal, aspectos sociais e ético (proteção à natureza, não violência, aquecimento social, sustentabilidade, ativismo), treinamentos realísticos ou com simulação, saúde e artes (expressão artísticas), e foram distribuídos conforme mostrados na Figura 1. Os jogos voltados para a área de saúde foram subdivididos em dois grupos: um de jogos voltados para a formação de profissionais e outro de jogos que têm por objetivo levar esclarecimentos para o público leigo.

A Tabela 4 mostra os métodos de avaliação de aprendizagem utilizados em cada área de aplicação dos jogos sérios citados nos artigos selecionados. No geral, em todas as áreas os métodos mais utilizados foram questionário e pré-teste com pós-teste. Uma exceção foi a área de Saúde com foco na formação de profissionais, que utilizou avaliação *in-game* e análise de *logs*, ao invés de questionários. Isso provavelmente se deve à neces-

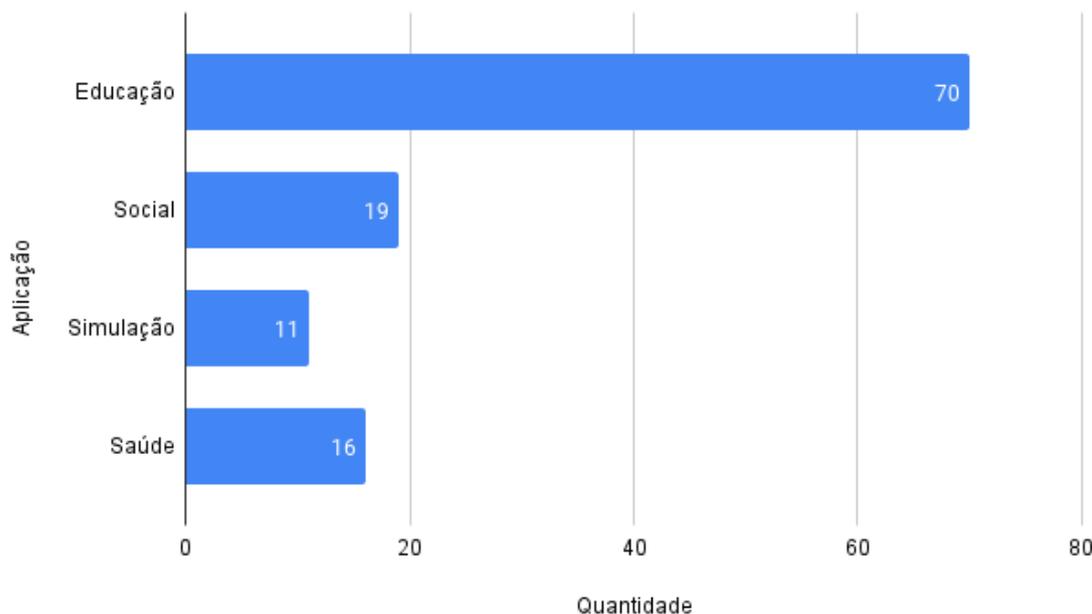


Figura 1. Área de aplicação dos jogos citados nos artigos selecionados

sidade de um acompanhamento mais próximo das ações do aprendiz, pela complexidade das tarefas a serem executadas.

5. Discussão

Ao analisar os métodos disponíveis, é crucial reconhecer que não há uma abordagem correta ou incorreta. O método mais adequado pode variar de acordo com diversos fatores, como a área de estudo da pesquisa e os recursos disponíveis para avaliar a aprendizagem. Portanto, é importante estar familiarizado com as práticas existentes e suas características.

As entrevistas oferecem uma abordagem mais personalizada, permitindo esclarecer dúvidas com os participantes e extrair informações relevantes. No entanto, é preciso ressaltar que esse método tende a limitar o número de participantes, devido ao tempo necessário para avaliar cada um individualmente. Essa restrição quantitativa pode afetar a representatividade dos resultados, já que uma amostra muito pequena pode enviesar e restringir a interpretação dos dados.

Por outro lado, os questionários consistem em uma série de perguntas padronizadas, o que facilita a coleta de dados de um grande número de participantes em um curto período de tempo. Embora os questionários possibilitem uma amostra maior, é importante reconhecer suas limitações. O formato limita a capacidade dos pesquisadores de obter informações mais detalhadas, podendo não capturar todas as nuances e aspectos relevantes para a pesquisa.

Apenas realizar uma avaliação ao final do experimento pode não ser suficiente, pois não permite afirmar que o conhecimento demonstrado no pós-teste foi adquirido exclusivamente por meio do jogo sério. Para mitigar esse efeito, é necessário realizar um pré-teste, que serve como um controle do conhecimento que os participantes já possuíam antes de se envolverem com o jogo.

Tabela 4. Área de aplicação do jogo e método de avaliação de aprendizagem

Método de avaliação	Educação	Simulação	Saúde	Social	Total geral
Atividade Prática	1				1
Avaliação in-game	2		2		4
Entrevista	1				1
Entrevista e análise de log			1		1
Fórum de discussões e log de usuário				1	1
Monitoramento das atividades	1				1
Pós-testes	4				4
Pré-teste e Pós-teste	28	5	8	5	46
Pré-teste, Pós-teste e entrevista				2	2
Pré-teste, Pós-teste e observação	1				1
Questionário e avaliação de especialista	1				1
Questionário e entrevista		2		1	3
Questionário	25	3	3	9	40
Questionário e Observação	1				1
Questionário e pré-teste e pós-teste	4	1	1	1	7
Questionário, avaliação in game	1				1
Questionário, observação e entrevista			1		1
Total geral	70	11	16	19	116

6. Conclusões

Neste trabalho, foram analisados os conteúdos de 116 artigos selecionados. Os métodos de avaliação de ganho de conhecimento por meio de jogos educativos foram principalmente pré-teste em conjunto com pós-teste e questionário, além de diversas combinações de métodos. O monitoramento, análise de *log* e avaliação *in-game* ainda são pouco utilizados. Como esperado, nos artigos estudados, quase todos tiveram participação do público-alvo na avaliação, exceto um em que o jogo foi desenvolvido para crianças com dislexia e contou com a participação de profissionais na avaliação.

Espera-se que a síntese destes métodos e práticas de avaliação sirvam como orientações para a comunidade da área de computação e desenvolvedores de jogos sérios, contribuindo assim para o avanço do uso desses jogos como ferramenta educacional.

Para trabalhos futuros, sugere-se uma análise mais detalhada dos artigos selecionados, abordando, por exemplo: local/país do estudo; se a avaliação estava vinculada à educação formal ou não-formal (independente da área), ou em tratamento clínico; a faixa etária do público-alvo do estudo. Percebe-se também, pelos resultados obtidos, a necessidade de implementar mais estratégias de avaliação *in-game*, incluindo o desenvolvimento de *frameworks* que possam facilitar a inclusão dessas estratégias em jogos educacionais.

Agradecimentos

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código Financeiro 001 e à PUC Minas pelo apoio financeiro.

Referências

Arawjo, I., Wang, C.-Y., Myers, A. C., Andersen, E., and Guimbretière, F. (2017). Teaching programming with gamified semantics. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pages 4911–4923.

- Chang, C.-C., Liang, C., Chou, P.-N., and Lin, G.-Y. (2017). Is game-based learning better in flow experience and various types of cognitive load than non-game-based learning? perspective from multimedia and media richness. *Computers in Human Behavior*, 71:218–227.
- Chittaro, L. (2016). Designing serious games for safety education: “learn to brace” versus traditional pictorials for aircraft passengers. *IEEE Transactions on Visualization and Computer Graphics*, 22(5):1527–1539.
- Chrysafiadi, K., Papadimitriou, S., and Virvou, M. (2019). Which is better for learning: a web-based educational application or an educational game? In *2019 International Symposium on Performance Evaluation of Computer and Telecommunication Systems (SPECTS)*, pages 1–6.
- Chrysafiadi, K., Papadimitriou, S., and Virvou, M. (2022). Cognitive-based adaptive scenarios in educational games using fuzzy reasoning. *Knowledge-Based Systems*, 250:109111.
- De Freitas, S. and Liarokapis, F. (2011). Serious games: a new paradigm for education? *Serious games and edutainment applications*, pages 9–23.
- DeSmet, A., Bastiaensens, S., Van Cleemput, K., Poels, K., Vandebosch, H., Deboutte, G., Herrewijn, L., Malliet, S., Pabian, S., Van Broeckhoven, F., De Troyer, O., Deglorie, G., Van Hoecke, S., Samyn, K., and De Bourdeaudhuij, I. (2018). The efficacy of the friendly attac serious digital game to promote prosocial bystander behavior in cyberbullying among young adolescents: A cluster-randomized controlled trial. *Computers in Human Behavior*, 78:336–347.
- Duin, H., Pourabdollahian, B., Thoben, K.-D., and Taisch, M. (2013). On the effectiveness of teaching sustainable global manufacturing with serious gaming. In *2013 International Conference on Engineering, Technology and Innovation (ICE) & IEEE International Technology Management Conference*, pages 1–8.
- Guo, Y. R. and Goh, D. H.-L. (2016). From storyboard to software: User evaluation of an information literacy game. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing, SAC '16*, page 199–201, New York, NY, USA. Association for Computing Machinery.
- Harpstead, E. and Alevan, V. (2015). Using empirical learning curve analysis to inform design in an educational game. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play, CHI PLAY '15*, page 197–207, New York, NY, USA. Association for Computing Machinery.
- Hooshyar, D., Malva, L., Yang, Y., Pedaste, M., Wang, M., and Lim, H. (2021). An adaptive educational computer game: Effects on students’ knowledge and learning attitude in computational thinking. *Computers in Human Behavior*, 114:106575.
- Ismail, S., Abdennadher, S., and Abouelsaadat, W. (2016). Rules on wheels: A serious game for teaching traffic signs. In *2016 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*, pages 1–7.
- Kechai, H. E. and Pierrot, L. (2015). Participatory design in eu-topia: A serious game for intercultural competences during work mobility. In *2015 IEEE 15th International Conference on Advanced Learning Technologies*, pages 127–131.

- Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., and Linkman, S. (2009). Systematic literature reviews in software engineering – a systematic literature review. *Information and Software Technology*, 51(1):7–15. Special Section - Most Cited Articles in 2002 and Regular Research Papers.
- Krath, J., SchÄ¼rmann, L., and von Korfflesch, H. F. (2021). Revealing the theoretical basis of gamification: A systematic review and analysis of theory in research on gamification, serious games and game-based learning. *Computers in Human Behavior*, 125:106963.
- Lee, T. Y., Mauriello, M. L., Ahn, J., and Bederson, B. B. (2014). Ctarcade: Computational thinking with games in school age children. *International Journal of Child-Computer Interaction*, 2(1):26–33.
- Liu, S. and Liu, M. (2020). The impact of learner metacognition and goal orientation on problem-solving in a serious game environment. *Computers in Human Behavior*, 102:151–165.
- Maertens, M., Vandewaetere, M., Cornillie, F., and Desmet, P. (2014). From pen-and-paper content to educational math game content for children: A transfer with added difficulty. *International Journal of Child-Computer Interaction*, 2(2):85–92. Special Issue: Learning from Failures in Game Design for Children.
- Michael, D. R. and Chen, S. L. (2005). *Serious games: Games that educate, train, and inform*. Muska & Lipman/Premier-Trade.
- Ouherrou, N., Elhammoumi, O., Benmarrakchi, F., and El Kafi, J. (2018). A heuristic evaluation of an educational game for children with dyslexia. In *2018 IEEE 5th International Congress on Information Science and Technology (CiSt)*, pages 386–390.
- Özgen, D. S., Afacan, Y., and Surer, E. (2020). Save the planets: A multipurpose serious game to raise environmental awareness and to initiate change. In *Proceedings of the 6th EAI International Conference on Smart Objects and Technologies for Social Good, GoodTechs '20*, page 132–137, New York, NY, USA. Association for Computing Machinery.
- Petri, G. and von Wangenheim, C. G. (2016). How to evaluate educational games: a systematic literature review. *Journal of Universal Computer Science*, 22:992–1021.
- Qin, J., Chui, Y.-P., Pang, W.-M., Choi, K.-S., and Heng, P.-A. (2010). Learning blood management in orthopedic surgery through gameplay. *IEEE Computer Graphics and Applications*, 30(2):45–57.
- Reynaldo, C., Christian, R., Hosea, H., and Gunawan, A. A. S. (2021). Using video games to improve capabilities in decision making and cognitive skill: A literature review. *Procedia Computer Science*, 179:211–221. 5th International Conference on Computer Science and Computational Intelligence 2020.
- Robertson, J., Macvean, A., Fawkner, S., Baker, G., and Jepson, R. G. (2018). Savouring our mistakes: Learning from the fitquest project. *International Journal of Child-Computer Interaction*, 16:55–67.
- Rodríguez-Cerezo, D., Sarasa-Cabezuelo, A., Gómez-Albarrán, M., and Sierra, J.-L. (2014). Serious games in tertiary education: A case study concerning the comprehen-

- sion of basic concepts in computer language implementation courses. *Computers in Human Behavior*, 31:558–570.
- Santos, A. A. and Pereira, O. J. (2020). A importância dos jogos e brincadeiras lúdicas na educação infantil. *Revista Eletrônica Pesquiseduca*, 11(25):480–493.
- Savi, R. and Ulbricht, V. R. (2008). Jogos digitais educacionais: Benefícios e desafios. *Revista Novas Tecnologias na Educação*, 6(1).
- Schoeffel, P., Wazlawick, R. S., and Ramos, V. F. C. (2018). Using multiple active teaching-learning approaches in software project management: A longitudinal analysis of students' motivation and learning. In *2018 IEEE Frontiers in Education Conference (FIE)*, pages 1–9.
- Sinclear, D., Birch Flensborg, L., Lindblad Fogsgaard, A., and Lochtefeld, M. (2022). Face-the-waste - learning about food waste through a serious game. In *20th International Conference on Mobile and Ubiquitous Multimedia, MUM 2021*, page 67–72, New York, NY, USA. Association for Computing Machinery.
- Ullah, M., Amin, S. U., Munsif, M., Safaev, U., Khan, H., Khan, S., and Ullah, H. (2022). Serious games in science education. a systematic literature review. *Virtual Reality & Intelligent Hardware*, 4(3):189–209. Advances in Wireless Sensor Networks under AI-SG for Augmented Reality Special Issue.
- von Wangenheim, C. G., Savi, R., and Borgatto, A. F. (2012). Deliver! – an educational game for teaching earned value management in computing courses. *Information and Software Technology*, 54(3):286–298.
- Wang, S.-T., Liu, L.-M., and Wang, S.-M. (2018). The design and evaluate of virtual reality immersive learning - the case of serious game “calcium looping for carbon capture”. In *2018 International Conference on System Science and Engineering (ICSSE)*, pages 1–4.
- Winter, M., Pryss, R., Probst, T., and Reichert, M. (2020). Learning to read by learning to write: Evaluation of a serious game to foster business process model comprehension. *JMIR Serious Games*, 8(1):e15374.
- Yasin, A., Liu, L., Li, T., Fatima, R., and Jianmin, W. (2019). Improving software security awareness using a serious game. *IET Software*, 13(2):159–169.