

Evolução de um Modelo de Avaliação de Jogos para o Ensino de Computação

Giani Petri^{1,2}, Christiane Gresse von Wangenheim¹, Adriano Ferretti Borgatto¹

¹Grupo de Qualidade de Software (GQS) – Programa de Pós-Graduação em Ciência da Computação (PPGCC) – Departamento de Informática e Estatística (INE) – Universidade Federal de Santa Catarina (UFSC) – Florianópolis – SC – Brasil

²Universidade Federal de Santa Maria (UFSM) – Santa Maria – RS – Brasil
gpetri@inf.ufsm.br, {c.wangenheim, adriano.borgatto}@ufsc.br

Abstract. *Educational games are assumed to be an efficient instructional strategy for computing education. However, it is essential to systematically evaluate such games in order to obtain sound evidence on their quality. A prominent model for the evaluation of educational games is MEEGA (Model for the Evaluation of Educational GAMES). However, analyzing the initial version we have identified limitations in terms of its validity and reliability. Therefore, the objective of this paper is to present an evolution of the model, MEEGA+. The MEEGA+ model has been systematically developed by decomposing evaluation goals into measures and defining measurement instruments to evaluate the perceived quality of educational games in terms of player experience and perceived learning from the viewpoint of the student and instructor. The model, thus, assists game creators, instructors and/or researchers to evaluate educational games and, hence, contribute to their improvement and adoption in practice.*

Resumo. *Jogos educacionais são considerados uma estratégia instrucional eficiente para o ensino de computação. No entanto, é essencial sistematicamente avaliar esses jogos a fim de obter evidências de sua qualidade. Um modelo de avaliação de jogos educacionais que têm se destacado é o MEEGA (Model for the Evaluation of Educational Games). No entanto, analisando a versão inicial identificou-se limitações em sua validade e confiabilidade. Assim, o objetivo deste trabalho é apresentar uma evolução do modelo, MEEGA+. O modelo MEEGA+ é sistematicamente desenvolvido decompondo o objetivo da avaliação em medidas e definindo instrumentos de medição para avaliar a qualidade de jogos em termos de experiência do jogador e percepção da aprendizagem do ponto de vista do aluno e instrutor. O modelo auxilia criadores de jogos, instrutores e/ou pesquisadores a avaliar jogos educacionais e, assim, contribuir para sua melhoria e adoção na prática.*

1. Introdução

Nos últimos anos, jogos educacionais estão sendo utilizados como uma estratégia instrucional inovadora para o ensino de computação [Calderón e Ruiz, 2015; Battistella

e Gresse von Wangenheim, 2016; Souza e França, 2016; Petri e Gresse von Wangenheim, 2017]. Esses jogos educacionais são projetados especificamente para ensinar as pessoas sobre um determinado assunto, expandir conceitos, reforçar o desenvolvimento, ou ajudá-los a explorar ou aprender uma habilidade ou uma mudança de atitude [Dempsey et al., 1996]. Motivados pela necessidade de fornecer mais oportunidades práticas para os estudantes de computação, uma grande variedade de jogos educacionais já foi desenvolvida [Battistella e Gresse von Wangenheim, 2016; Petri e Gresse von Wangenheim, 2017]. A maioria dos jogos desenvolvidos são digitais, com uma tendência também em jogos não-digitais (de tabuleiros, cartas, etc.) para o ensino de diversas áreas de conhecimento da computação, com ênfase nas áreas de engenharia de software, fundamentos de desenvolvimento de software e algoritmos [Battistella e Gresse von Wangenheim, 2016; Petri e Gresse von Wangenheim, 2017].

Assim, considera-se que os jogos educacionais possam ser uma estratégia instrucional eficaz e eficiente para o ensino de computação. No entanto, estas alegações são questionáveis ou não rigorosamente comprovadas [Gresse von Wangenheim e Shull, 2009; Caulfield et al., 2011; Petri e Gresse von Wangenheim, 2017]. Na prática, os jogos para o ensino de computação necessitam mostrar o impacto esperado de aprendizagem e/ou o engajamento que eles prometem [Caulfield et al., 2011]. Portanto, é essencial avaliar sistematicamente a qualidade desses jogos, a fim de obter provas sólidas sobre a sua qualidade [Gresse von Wangenheim e Shull, 2009; Caulfield et al., 2011; Petri e Gresse von Wangenheim, 2017]. Neste contexto, foram realizados alguns trabalhos para obter evidências empíricas da qualidade e do impacto dos jogos na aprendizagem dos alunos [Gresse von Wangenheim e Shull, 2009; Caulfield et al., 2011; Calderón e Ruiz, 2015; Petri e Gresse von Wangenheim, 2017]. No entanto, os resultados dos trabalhos mostram que a maioria das avaliações de jogos educacionais são realizados de forma *ad-hoc* em termos de design de pesquisa, medição, coleta de dados e análise, evidenciando assim a falta de rigor científico nas avaliações [Calderón e Ruiz, 2015; Petri e Gresse von Wangenheim, 2017].

Neste contexto, para fornecer um apoio mais sistemático, até agora houve poucas tentativas de desenvolver abordagens para avaliar a qualidade de jogos educacionais [Petri e Gresse von Wangenheim, 2016]. Entre os modelos utilizados para a avaliação de jogos, o modelo MEEGA (*Model for the Evaluation of Educational Games*) [Savi et al., 2011] é identificado como um dos mais utilizados na prática [Calderón e Ruiz, 2015; Petri e Gresse von Wangenheim, 2016; Petri e Gresse von Wangenheim, 2017]. MEEGA é um modelo desenvolvido para a avaliação de jogos educacionais por meio da aplicação de um questionário padronizado após os estudantes utilizarem o jogo. O modelo é desenvolvido usando a abordagem GQM (*Goal/Question/Metric*) [Basili et al., 1994] para definir explicitamente um programa de medição para avaliar os jogos em três fatores: motivação, experiência do usuário e aprendizagem do ponto de vista do aluno. A partir dos fatores definidos é desenvolvido um questionário para a coleta de dados. No entanto, analisando a confiabilidade e validade, identificou-se que o modelo MEEGA apresenta algumas limitações em sua concepção e avaliação. As limitações são evidenciadas no resultado da análise de sua validade que indica uma decomposição diferente da estrutura original do MEEGA, com sobreposição de alguns conceitos entre os fatores de motivação e experiência do usuário,

além da falta de compreensão dos alunos de alguns itens do instrumento de medição [Petri et al., 2017].

Diante disso, o objetivo deste trabalho é apresentar o modelo MEEGA+, uma evolução do modelo MEEGA.

2. Metodologia

Com o objetivo de evoluir a versão inicial do modelo MEEGA, é adotada uma metodologia de pesquisa multimétodo, conforme apresentado na Figura 1.

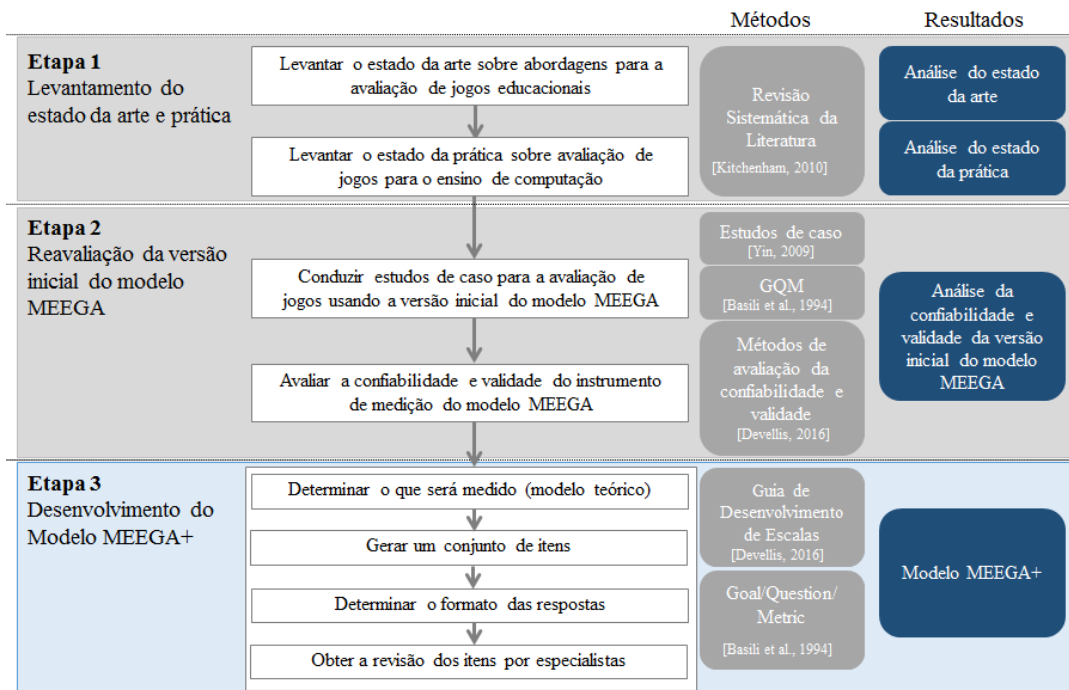


Figura 1. Metodologia de pesquisa

Etapa 1 - Levantamento do Estado da Arte e Prática. Inicialmente, foi realizado um levantamento do estado da arte sobre as abordagens existentes para sistematicamente avaliar os jogos educacionais. Como resultado, identificou-se que há poucas abordagens para avaliar jogos. Além disso, a maioria das abordagens não fornecem uma explícita definição de um objetivo de avaliação, medidas ou instrumento de coleta de dados [Petri e Gresse von Wangenheim, 2016]. Adicionalmente, também foi realizado um levantamento do estado da prática sobre como os jogos para o ensino de computação são avaliados. Como resultado observou-se que a maioria das avaliações são realizadas de maneira *ad-hoc* em termos de design de pesquisa, medição, coleta e análise de dados, evidenciando a falta de rigor científico nas avaliações realizadas [Petri e Gresse von Wangenheim, 2017]. Os resultados e a descrição detalhada de cada uma das revisões da literatura podem ser encontrados em [Petri e Gresse von Wangenheim, 2016; Petri e Gresse von Wangenheim, 2017].

Etapa 2 – Reavaliação da versão inicial do Modelo MEEGA. De modo a reavaliar em larga-escala a versão inicial do modelo MEEGA, foi realizado um estudo de caso. O estudo de caso analisou 1000 dados coletados de 60 estudos que reportaram a avaliação de um jogo usando o modelo MEEGA. Com base nos dados foi analisado o modelo MEEGA em termos de confiabilidade e validade. Os resultados indicam que o

questionário do modelo MEEGA é aceitável em termos de confiabilidade (alfa de Cronbach = 0,915). No entanto, em termos de validade, a análise fatorial indicou uma decomposição diferente da estrutura original do MEEGA, que é composto por três fatores (motivação, experiência do usuário e aprendizagem). Isto indica que alguns conceitos entre os fatores de motivação e experiência do usuário se sobrepõem e então, precisam ser reagrupados. A descrição detalhada do estudo de caso e da análise podem ser encontrados em [Petri et al., 2017].

Etapas 3 – Desenvolvimento do Modelo MEEGA+. Nesta etapa o modelo MEEGA+ é desenvolvido, levando em consideração o levantamento do estado da arte (Etapa 1) e os resultados da reavaliação do modelo MEEGA (Etapa 2). O desenvolvimento do modelo MEEGA+ segue as etapas do guia de desenvolvimento de escalas proposto por DeVellis (2016):

Determinar o que será medido (modelo teórico). De modo a sistematicamente definir um programa de medição a abordagem GQM [Basili et al., 1994] é utilizada. Com base na abordagem GQM o objetivo da avaliação é definido, bem como a decomposição sistemática do objetivo da avaliação em fatores de qualidade a serem medidos (modelo teórico) e em itens do instrumento de medição.

Gerar um conjunto de itens. Com base no modelo teórico definido, adotando a abordagem GQM, os fatores do modelo teórico são sistematicamente decompostos em itens de questionário. Deste modo, o conjunto dos itens do instrumento de medição estão relacionados a um fator a ser medido e alinhados ao modelo teórico.

Determinar o formato das respostas. Nesta etapa, o formato de resposta para os itens do instrumento de medição é definido.

Obter a revisão dos itens por especialistas. Esta etapa objetiva analisar a validade aparente do instrumento de medição do ponto de vista de especialistas.

As etapas de administrar os itens para um grupo de indivíduos e análise de desempenho do instrumento de medição estão em fase de execução e estão fora do escopo deste artigo.

3. O Modelo MEEGA+

Esta seção apresenta o modelo MEEGA+, uma evolução do modelo MEEGA voltado para a avaliação da qualidade de jogos educacionais para o ensino de computação. Por falta de uma definição comumente aceita sobre qualidade de jogos educacionais, neste trabalho, entende-se que um jogo de qualidade possui os objetivos educacionais bem definidos, promove a aprendizagem de conteúdos e uma agradável experiência aos jogadores.

O objetivo do modelo MEEGA+ é: analisar *jogos educacionais* com o propósito de avaliar a *percepção da qualidade* em termos de *experiência do jogador e percepção da aprendizagem* do ponto de vista de *alunos e instrutores* no contexto de *curso superiores da área de computação*. A partir do objetivo de avaliação, são derivadas as medidas referentes aos dados a serem coletados durante o estudo para que se atinja o objetivo da avaliação. Esta definição é feita por meio de uma decomposição hierárquica, realizada com base na revisão da literatura e na avaliação da versão inicial do modelo MEEGA.

O fator de qualidade referente a experiência do jogador é definido com base na revisão da literatura (Etapa 1 da metodologia). Embora observa-se a falta de um padrão consistente dos fatores de qualidade a serem avaliados, há uma tendência que abrange um conjunto de fatores de qualidade relacionados principalmente à motivação, experiência do usuário, usabilidade, engajamento e apreciação (*enjoyment*) [Petri e Gresse von Wangenheim, 2017]. Esses fatores de qualidade ainda estão fragmentados em dimensões que se sobrepõem conceitualmente, como por exemplo, os fatores imersão e atenção focada. Além disso, os resultados obtidos na análise fatorial na avaliação da versão inicial do modelo MEEGA mostram que os conceitos de motivação e experiência do usuário se sobrepõem. Assim, é realizado um mapeamento para identificar as dimensões semelhantes. Os detalhes do mapeamento estão apresentados em [Petri et al., 2016]. A partir do mapeamento, define-se para o modelo MEEGA+ o fator de qualidade de experiência do jogador composto por um conjunto de dimensões: atenção focada, diversão, desafio, interação social, confiança, relevância, satisfação e usabilidade.

A usabilidade ainda está fragmentada em cinco subdimensões: aprendizibilidade, operabilidade, estética, acessibilidade e proteção de erro do usuário. Em relação a acessibilidade (subdimensão da usabilidade), o modelo MEEGA+ se concentra em avaliar se o jogo é acessível por pessoas com deficiência visual baixa/moderada e/ou com daltonismo. Esta decisão é tomada com base que o número de pessoas com deficiência visual baixa/moderada representa 15,1% da população brasileira [IBGE, 2012]. A subdimensão de proteção de erro de usuário é usada exclusivamente para avaliar jogos educacionais digitais. Assim, ao avaliar jogos não-digitais esta subdimensão deve ser desconsiderada.

O fator de qualidade em relação à aprendizagem percebida é refinado em duas dimensões. A avaliação da aprendizagem de curto prazo baseia-se no modelo de avaliação de Sindre e Moody (2003). O objetivo desta dimensão é avaliar o efeito global do jogo sobre a aprendizagem dos alunos no curso [Sindre e Moody, 2003]. A avaliação da dimensão dos objetivos de aprendizagem é personalizada de acordo com os objetivos de aprendizagem de cada jogo educacional. Tipicamente, os jogos para o ensino de computação são usados para melhorar o conhecimento sobre os níveis cognitivos de lembrança, compreensão e aplicação [ACM/IEEE-CS, 2013] de acordo com a versão revisada da taxonomia de Bloom [Anderson et al., 2001]. Estes três níveis de aprendizagem são considerados, pois representam o conhecimento que pode ser aprendido durante um curso de graduação [ACM/IEEE-CS, 2013]. No entanto, o modelo MEEGA+ também é flexível para cobrir os objetivos de avaliação em outros níveis cognitivos tais como: análise, avaliação e criação. Além do conhecimento técnico, os jogos também podem contribuir para o desenvolvimento de um conjunto de habilidades, tais como a resolução de problemas, comunicação interpessoal, trabalho em equipe, liderança, entre outros [ACM/IEEE-CS, 2013]. Tais objetivos de aprendizagem visam a mudança e/ou o desenvolvimento de comportamentos ou habilidades. Eles podem ser classificados em níveis de acordo com a taxonomia de Bloom do domínio psicomotor [Simpson, 1972]. Além disso, os jogos podem contribuir também para o desenvolvimento de atitudes profissionais, tais como o comportamento do estudante, considerando as questões profissionais, bem como uma atitude ética na profissão [ACM/IEEE-CS, 2013]. Objetivos de aprendizagem são tipicamente relacionados a

melhoria de atitudes, emoções e sentimentos e podem ser classificados nos níveis da taxonomia de Bloom do domínio afetivo [Kratwohl et al., 1973].

Desta forma, a decomposição hierárquica dos fatores de qualidade em dimensões/subdimensões resulta no modelo teórico (Figura 2).

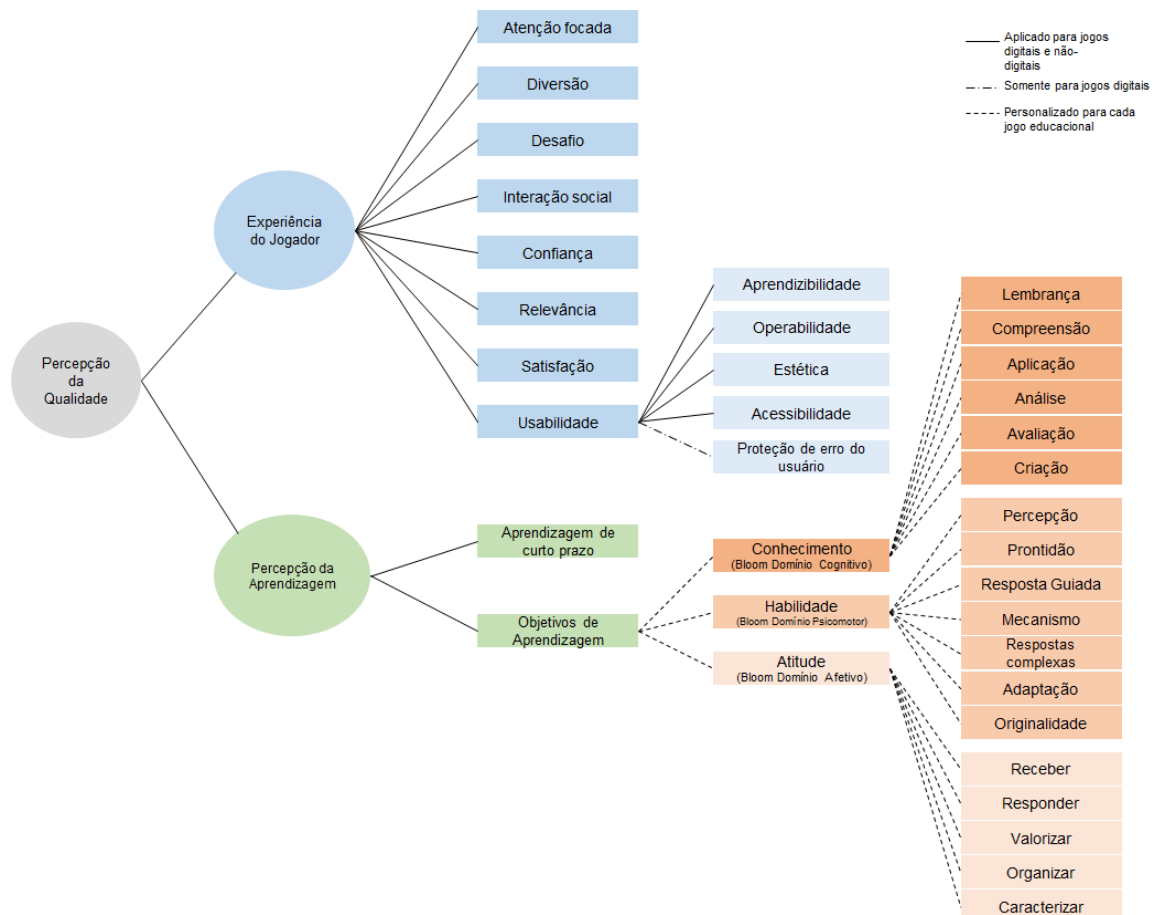


Figura 2. Decomposição dos fatores de qualidade

Hipótese e estratégia de pesquisa. A hipótese é que o jogo educacional contribui positivamente para a unidade instrucional, atinge seus objetivos de aprendizagem e promove uma agradável experiência ao jogador. A estratégia de pesquisa é selecionada com base no objetivo da avaliação e nas restrições práticas típicas das unidades instrucionais de ensino de computação. Isto significa que deve ser possível realizar a avaliação de forma rápida e não intrusiva, minimizando a interrupção do fluxo normal da aula. Portanto, opta-se pela realização de estudos de caso que permite uma pesquisa aprofundada de um indivíduo, grupo ou evento [Yin, 2009]. O design de pesquisa é definido como não-experimental com pós-teste (*one-shot post-test only design*) com um único grupo. Neste tipo de pesquisa, o estudo de caso começa com a aplicação do tratamento (jogo educacional) e, em seguida, um questionário é respondido pelos alunos para a coleta dos dados. Adotando esta estratégia de pesquisa, o objetivo da avaliação é avaliado com base nas percepções dos alunos (qualidade percebida) por meio de um questionário respondido após a aplicação do jogo em acordo com o objetivo de avaliação no nível 1 do modelo de Kirkpatrick [Kirkpatrick e Kirkpatrick, 2006].

Desenvolveu-se um instrumento de medição seguindo o método proposto por DeVellis (2016). O questionário é desenvolvido com base no modelo teórico (Figura 2), personalizando e unificando os questionários padrões existentes na literatura [Keller, 1987; Tullis e Albert, 2008; Sindre e Moody, 2003; Sweetser e Wyeth, 2005; Poels et al. 2007; Gámez, 2009; Takatalo et al., 2010; O'Brien e Toms, 2010; Wiebe et al., 2014; Sweetser e Wyeth, 2005; Fu, Su, e Yu, 2010; Brooke, 1996; Davis, 1989]. A descrição detalhada do questionário está disponível em [Petri et al., 2016]. O formato de resposta para cada um dos itens do instrumento de medição é baseado em uma escala de Likert de 5 pontos com alternativas de resposta que variam de discordo totalmente a concordo totalmente [DeVellis, 2016]. A parte inicial do questionário do modelo MEEGA+ objetiva coletar informações demográficas para identificar o perfil dos alunos. Os itens do questionário que objetivam coletar dados sobre o fator de qualidade de experiência do jogador, quando da avaliação de jogos não-digitais do ponto de visto do aluno, são apresentados na Tabela 1.

Tabela 1. Itens do questionário para a avaliação da experiência do jogador (jogos não-digitais)

No.	Dimensão/Subdimensão	Descrição do Item
1	Atenção focada	Houve algo interessante no início do jogo que capturou minha atenção.
2		Eu estava tão envolvido no jogo que eu perdi a noção do tempo.
3		Eu esqueci sobre o ambiente ao meu redor enquanto jogava este jogo.
4	Diversão	Eu me diverti com o jogo.
5		Aconteceu alguma situação durante o jogo (elementos do jogo, competição, etc.) que me fez sorrir.
6	Desafio	Este jogo é adequadamente desafiador para mim.
7		O jogo oferece novos desafios (oferece novos obstáculos, situações ou variações) com um ritmo adequado.
8		O jogo não se torna monótono nas suas tarefas (repetitivo ou com tarefas chatas).
9	Interação social	Eu pude interagir com outras pessoas durante o jogo.
10		O jogo promove momentos de cooperação e/ou competição entre os jogadores.
11		Eu me senti bem interagindo com outras pessoas durante o jogo.
12	Confiança	Quando olhei pela primeira vez o jogo, eu tive a impressão de que seria fácil para mim.
13		A organização do conteúdo me ajudou a estar confiante de que eu iria aprender com este jogo.
14	Relevância	O conteúdo do jogo é relevante para os meus interesses.
15		É claro para mim como o conteúdo do jogo está relacionado com a disciplina.
16		O jogo é um método de ensino adequado para esta disciplina.
17		Eu prefiro aprender com este jogo do que de outra forma (outro método de ensino).
18	Satisfação	Completar as tarefas do jogo me deu um sentimento de realização.
19		É devido ao meu esforço pessoal que eu consigo avançar no jogo.
20		Me sinto satisfeito com as coisas que aprendi no jogo.
21		Eu recomendaria este jogo para meus colegas.
22	Estética	O design do jogo é atraente (interface, gráficos, tabuleiro, cartas, etc.).
23		Os textos, cores e fontes combinam e são consistentes.
24		Eu precisei aprender poucas coisas para poder começar a jogar o jogo.
25	Aprendizabilidade	Aprender a jogar este jogo foi fácil para mim.
26		Eu acho que a maioria das pessoas aprenderiam a jogar este jogo rapidamente.
27	Operabilidade	Eu considero que o jogo é fácil de jogar.
28		As regras do jogo são claras e compreensíveis.
29	Acessibilidade	As fontes (tamanho e estilo) utilizadas no jogo são legíveis.
30		As cores utilizadas no jogo são compreensíveis.

Adicionalmente, quando o estudo envolve a avaliação de jogos digitais, três itens relacionados a usabilidade devem ser adicionados ao questionário, conforme apresentados na Tabela 2.

Tabela 2. Itens adicionais para a avaliação da experiência do jogador (jogos digitais)

No.	Dimensão/Subdimensão	Descrição do Item
31	Usabilidade	Acessibilidade
32		Proteção de erro do usuário
33		Quando eu cometo um erro é fácil de me recuperar rapidamente.

A avaliação da percepção da aprendizagem é dividida em duas dimensões (aprendizagem de curto prazo e objetivos de aprendizagem). Os itens do questionário

referentes a aprendizagem de curto prazo são padrões e utilizados para avaliar todos os jogos (digitais e não-digitais). Os itens referentes a subdimensão dos objetivos de aprendizagem devem ser customizados conforme os objetivos definidos para cada jogo educacional. No entanto, o questionário oferece um *template* de como customizar os itens do questionário, conforme o nível em que os objetivos de aprendizagem do jogo estão definidos, considerando os níveis cognitivo, psicomotor e afetivo da taxonomia de Bloom. Cada objetivo de aprendizagem do jogo deve corresponder a exatamente um item do questionário, conforme exemplo apresentado na Tabela 3.

Tabela 3. Itens do questionário para a avaliação da percepção da aprendizagem

No.	Dimensão/Subdimensão	Descrição do Item
34	Percepção da Aprendizagem	O jogo contribuiu para a minha aprendizagem na disciplina.
35		O jogo foi eficiente para minha aprendizagem, em comparação com outras atividades da disciplina.
36		O jogo contribuiu para <verbo conforme nível do objetivo de aprendizagem (cognitivo, psicomotor, afetivo)> <objetivo/conceito>.
...		Um exemplo conforme o objetivo de aprendizagem do jogo SCRUMIA (Gresse von Wangenheim et al., 2013): O jogo contribuiu para <i>relembrar</i> os conceitos sobre o <i>Planejamento de uma Sprint</i> .

4. Avaliação inicial do Modelo MEEGA+

Com o objetivo de avaliar a validade aparente do modelo MEEGA+ foi realizado uma avaliação inicial por meio de um painel de especialistas. O painel de especialistas é composto por um grupo multidisciplinar de especialistas (pesquisadores seniores) das áreas de computação e estatística. A revisão dos especialistas objetiva analisar a clareza, relevância, consistência e completude dos fatores/dimensões do modelo e dos itens do instrumento de medição do modelo MEEGA+.

Uma reunião foi agendada com cada especialista de modo a apresentar e explicar o desenvolvimento do modelo MEEGA+ e do instrumento de medição. Para facilitar a análise, cada especialista também recebeu uma cópia impressa da decomposição dos fatores do modelo (Figura 2), bem como, do instrumento de medição. Na sequência, os especialistas analisaram e avaliaram o modelo MEEGA+ e o instrumento de medição.

De modo geral, os especialistas confirmaram a relevância do modelo MEEGA+ para a área de pesquisa, pelo fato de, atualmente, existir uma carência no rigor científico aplicado na avaliação de jogos educacionais. De mesmo modo, os especialistas também confirmaram a completude, indicando que o modelo representa os fatores importantes para a avaliação da qualidade dos jogos educacionais. Além disso, os especialistas também indicaram a clareza e a consistência na decomposição dos fatores que teve como base a revisão da literatura e a análise da versão inicial do modelo, bem como, a descrição dos itens do instrumento de medição. No entanto, algumas alterações na redação e na formatação do texto dos itens foram sugeridas pelos especialistas. Todas as sugestões de alteração dos itens foram consideradas na versão final do instrumento de medição do modelo MEEGA+.

5. Conclusão

Este artigo apresenta o desenvolvimento do modelo MEEGA+, uma evolução do modelo MEEGA, atualmente o modelo de avaliação de jogos educacionais mais utilizado na prática. O modelo MEEGA+ é sistematicamente desenvolvido decompondo o objetivo da avaliação em medidas e definindo um instrumento de medição para avaliar

a percepção da qualidade de jogos educacionais em termos de experiência do jogador e percepção da aprendizagem com base na literatura e experiências práticas. Como próxima etapa desta pesquisa, está sendo realizado diversos estudos de caso para a avaliação de vários jogos educacionais em diferentes cursos/instituições de ensino, de modo a coletar dados para avaliar a confiabilidade e validade do modelo MEEGA+.

Agradecimentos

Este trabalho é apoiado pelo CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), uma entidade do governo brasileiro focada no desenvolvimento científico.

Referências

- ACM/IEEE-CS. (2013). Computer Science Curricula 2013: Curriculum Guidelines for Undergraduate Degree Programs in Computer Science. Disponível em: <<https://www.acm.org/education/CS2013-final-report.pdf>> Acesso: 17 outubro 2016.
- Anderson, L. W., Krathwohl, D. R., e Bloom, B. S. (2001). A taxonomy for learning, teaching, and assessing: a revision of Bloom's taxonomy of educational objectives. Longman.
- Basili, V. R., Caldiera, G., e Rombach, H. D. (1994). Goal, Question Metric Paradigm. In J. J. Marciniak, Encyclopedia of Software Engineering. New York, USA.
- Battistella, P. e Wangenheim, C. G. (2016). Games for Teaching Computing in Higher Education – A Systematic Review. In *IEEE Technology and Engineering Education Journal*, 9(1), 8-30.
- Brooke, J. (1996). SUS-A quick and dirty usability scale. In *Usability evaluation in industry*, 189(194), 4-7.
- Calderón A. e Ruiz M. (2015). A systematic literature review on serious games evaluation: An application to software project management. In *Computers & Education*, 87, 396-422.
- Caulfield, C., Xia, J., Veal, D., e Maj, S. P. (2011). A systematic survey of games used for software engineering education. In *Modern Applied Science*, 5(6), 28-43.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, p. 319-340.
- Dempsey, J., Rasmussen, K., e Lucassen, B. (1996). The instructional gaming literature: Implications and 99 sources. Technical Report 96-1. College of Education, University of South Alabama, USA.
- DeVellis, R. F. (2016). Scale development: theory and applications. 4. ed. SAGE Publications.
- Fu, F., Su, R., e Yu, S. (2009). EGameFlow: A scale to measure learners' enjoyment of e-learning games. In *Computers & Education*, 52(1), 101-112.
- Gámez, E. H. (2009). On the Core Elements of the Experience of Playing Video Games (Dissertation). UCL Interaction Centre, Department of Computer Science, London, UK.
- Gresse von Wangenheim, C. e Shull, F. (2009). To Game or Not to Game? In *IEEE Software*, 26(2), 92-94.
- Gresse von Wangenheim, C., Savi, R., e Borgatto, A. F. (2013). SCRUMIA - An educational game for teaching SCRUM in computing courses. In *Journal of Systems and Software*, 86(10), 2675-2687.
- Instituto Brasileiro de Geografia e Estatística (IBGE). (2012). Censo Demográfico 2010: Características gerais da população, religião e pessoas com deficiência.

- Keller, J. M. (1987). Development and Use of the ARCS Model of motivational Design. In *Journal of Instructional Development*, 10(3), 2-10.
- Kirkpatrick, D. L. e Kirkpatrick, J. D. (2006). Evaluating training programs: the four levels, 3. ed. Berrett-Koehler Publishers, USA.
- Kitchenham, B. (2010). Systematic literature reviews in software engineering – A tertiary study. *Information and Software Technology*, v. 52, n. 1, p. 792-805.
- Krathwohl, D. R.; Bloom, B. S.; e Masia, B. B. (1973). Taxonomy of Educational Objectives, the Classification of Educational Goals. Handbook II: Affective Domain. Nova York, EUA, 1973.
- O'Brien, H. L. e Toms, E. G. (2010). The Development and Evaluation of a Survey to Measure User Engagement. In *Journal of the American Society for Information Science and Technology*, 61(1), 50–69.
- Petri, G. e Gresse von Wangenheim, C. (2016). How to Evaluate Educational Games: a Systematic Literature Review. In *Journal of Universal Computer Science*, 22(7), 992-1021.
- Petri, G. e Gresse von Wangenheim, C. (2017). How games for computing education are evaluated: a systematic literature review. In *Computers & Education*, 107, 68-90.
- Petri, G.; Gresse von Wangenheim, C., e Borgatto, A. F. (2016). MEEGA+: An Evolution of a Model of Educational Games. Relatório Técnico INCoD/GQS.03.2016.E. INCoD/INE/UFSC, Florianópolis, Brazil.
- Petri, G.; Gresse von Wangenheim, C., e Borgatto, A. F. (2017). A large-scale evaluation of a model for the evaluation of games for teaching software engineering. Proc. of the 39th Int. Conf. on Software Engineering. Buenos Aires, Argentina.
- Poels, K., Kort, Y. D., e Ijsselstein, W. (2007). It is always a lot of fun!: exploring dimensions of digital game experience using focus group methodology. Proc. of Conf. on Future Play, Toronto, Canada.
- Savi, R., Gresse von Wangenheim, C., e Borgatto, A. F. (2011). A Model for the Evaluation of Educational Games for Teaching Software Engineering. Proc. of the 25th Simpósio Brasileiro de Engenharia de Software, São Paulo, Brazil.
- Sindre, G. e Moody, D. (2003). Evaluating the Effectiveness of Learning Interventions: an Information Systems Case Study. Proc. of the 11th European Conf. on Information Systems, Naples, Italy.
- Souza, M. e França, C. (2016). O que Explica o Sucesso de Jogos no Ensino de Engenharia de Software? Uma Teoria de Motivação. 24º Workshop sobre Educação em Computação, 2255-2264. Porto Alegre/RS, Brasil.
- Sweetser, P. e Wyeth, P. (2005). GameFlow: a model for evaluating player enjoyment in games. In *Computers in Entertainment*, 3(3), 1-24.
- Takatalo, J., Häkkinen, J., Kaistinen, J., e Nyman, G. (2010). Presence, Involvement, and Flow in Digital Games. In: Bernhaupt, R. (Ed.). *Evaluating User Experience in Games: Concepts and Methods*, Springer.
- Tullis, T. e Albert, W. (2008). *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*. Morgan Kaufmann.
- Wiebe, E. N., Lamb, A., Hardy, M., e Sharek, D. (2014). Measuring engagement in video game-based environments: Investigation of the User Engagement Scale. In *Computers in Human Behavior*, 32, 123–132.
- Yin, R.K. (2009). *Case study research: design and methods*, 4th ed. Sage Publications, USA.