

Taxonomia de Critérios para Avaliação de *Software* Educativo – TaCASE

Ozonias de O. B. Junior¹, Yuska P. C. Aguiar²

¹Centro de Informática – Universidade Federal de Pernambuco (UFPE)
Caixa Postal 7851 – 50.740-560 – Recife, PE – Brasil

²Departamento de Ciências Exatas – Universidade Federal da Paraíba (UFPB) – Rio Tinto, PB – Brasil

oob@cin.ufpe.br, yuska@dcx.ufpb.br

Abstract. *The diversity of approaches to evaluation of Educational Software (SE) is a reality. In order to promote a uniform understanding of the concepts used in this process, a taxonomy of evaluation criteria is useful. In view of this objective, a taxonomy of criteria for SE evaluation was conceived from 14 approaches described in the literature. This includes aspects of quality of use and software. In order to verify the adequacy of the mentioned taxonomy with the potential evaluators of SE, 48 specialists from the areas of Usability Engineering and Software Engineering answered a questionnaire elaborated in the scope of this research. The feedback obtained allowed to make refinements in the initial proposition.*

Resumo. *A diversidade de abordagens para avaliação de Software Educativo (SE) é uma realidade. Para favorecer a compreensão uniforme dos conceitos utilizados nesse processo, uma taxonomia de critérios de avaliação se mostra útil. Tendo em vista este objetivo, uma taxonomia de critérios para avaliação de SE foi concebida a partir de 14 abordagens descritas na literatura. Esta contempla aspectos de qualidade de uso e de software. A fim de verificar a adequação da referida taxonomia junto aos potenciais avaliadores de SE, 48 especialistas das áreas de Engenharia da Usabilidade e Engenharia de Software responderam a um questionário elaborado no escopo desta pesquisa. O retorno obtido permitiu realizar refinamentos na proposição inicial.*

1. Introdução

A combinação das áreas de Educação e Tecnologia da Informação e Comunicação (TIC) favorece a elevação do nível da qualidade do processo de ensino-aprendizagem, afetando diretamente os métodos, práticas e recursos de ensino [Coomans et al., 2015]. Segundo Peixoto, Resende e Pádua (2015), recursos educacionais digitais expandem o potencial e a eficácia do processo de ensino-aprendizagem tradicional, uma vez que a utilização correta desses recursos propicia o aprendizado lúdico do estudante, melhora a interpretação do conteúdo teórico a partir de sua aplicação prática, exercita a resolução de problemas em tempo ágil, possibilita que os discentes aprendam com seus erros, reduz a passividade dos educandos e, conseqüentemente, enriquece a prática pedagógica com a utilização de recursos multimídia.

O interesse da comunidade científica pelo desenvolvimento, pela avaliação e pelo uso de *Softwares* Educativos (SE) fica evidente diante das publicações em torno do

tema [De França et al., 2014]. Neste contexto, vale destacar que para alcançar os benefícios supracitados, faz-se necessário que os recursos educacionais digitais (i) tenham qualidade de acordo com os preceitos da Engenharia de *Software* e (ii) que seu projeto tenha sido direcionado às características, necessidades e habilidades de seus usuários finais, tal como recomenda a Engenharia da Usabilidade. A incidência de inadequações nessas vertentes pode comprometer seu objetivo de potencializar o processo de ensino-aprendizagem. Portanto, a adoção desses recursos deve ser precedida por uma avaliação referente à sua adequação em relação ao contexto em que se pretende aplicar [De França et al., 2014].

Para que a avaliação do SE seja eficiente é necessário realizar uma análise criteriosa considerando diferentes aspectos [Brito Junior e Aguiar, 2014], por exemplo: adequação, acurácia, maturidade, tolerância a falhas, estabilidade, prevenção de erros, ajuda e documentação, adaptabilidade, compatibilidade, outros. No entanto, diante da diversidade de abordagens existentes para avaliação de SE, alguns desafios persistem. Segundo Pereira et al. (2016a) alguns destes são: (i) grande variedade de abordagens, cada uma com critérios próprios e notações diferenciadas; (ii) algumas abordagens possuem critérios muito subjetivos e com alto nível de abstração, não apresentando explicações sobre como estes devem ser avaliados; (iii) os resultados alcançados em uma avaliação são fortemente vinculados ao avaliador que aplicou a abordagem; (iv) avaliações de SE geralmente contemplam critérios pedagógicos, de usabilidade e de qualidade de *software* (desenvolvimento e programação), o que exige multidisciplinaridade e conhecimentos específicos dos avaliadores, impactando na curva de aprendizado para realização das avaliações etc.

Além disso, Da Silva et al. (2016) sugerem que as abordagens existentes na literatura são complementares e que uma combinação destas deve ser aplicada para que a avaliação seja abrangente e considere os diferentes aspectos relevantes de um SE. Esta prática implica na necessidade de conhecer abordagens complementares para conduzir a avaliação. Segundo Pereira et al. (2016b), este cenário demonstra a importância de se estabelecer um estudo de equivalência entre as proposições existentes na literatura de forma a tornar a combinação destas homogênea. Ou seja, a concepção de uma Taxonomia de Critérios para Avaliação de Software Educativo (TaCASE) se mostra útil uma vez que favorece a compreensão uniforme dos conceitos utilizados nesse processo.

A Taxonomia de Critérios para Avaliação de *Software* Educativo concebida no contexto desta pesquisa contempla a combinação de um conjunto de 14 abordagens para avaliação de SE, previamente identificadas em revisão de literatura [Brito Junior et al., 2016a]. Com objetivo de diminuir a curva de aprendizagem necessária para a avaliação do SE, TaCASE se apoia nas nomenclaturas e conceitos amplamente utilizados na Engenharia de *Software* e na Engenharia de Usabilidade a partir de modelos de qualidade, normas e heurísticas adotadas nessas áreas de conhecimento [Brito Junior et al., 2016b]. Tendo conhecimento sobre as abordagens e suas características foi realizada a definição do conjunto de termos (critérios) para cada modelo de qualidade proposto. Com a seleção dos termos (critérios) finalizada para cada modelo de qualidade foi executado um mapeamento minucioso entre os critérios dos modelos de qualidade e os critérios das 14 abordagens. Nesse sentido, o presente trabalho busca apresentar o processo de concepção e validação da TaCASE proposta.

2. Taxonomia: conceitos e aplicações

A taxonomia é a ciência da identificação, utilizada para designar o conjunto de termos representativos de uma área, estruturados hierarquicamente. Esses termos são elementos estruturantes, estratégicos e centrais, utilizados para nomear, classificar e organizar entidades em grupos que compartilham características similares [Marques et al., 2004]. Na Ciência da Informação, taxonomia é um sistema utilizado para classificar e facilitar o acesso à informação, sendo seu objetivo representar conceitos através de termos, melhorar a comunicação entre especialistas e outros públicos, propor formas de controle da diversificação e oferecer um mapa do processo de conhecimento [Brasil et al., 2013].

O trabalho de Ziegler et al. (2004) apresenta uma taxonomia que faz a classificação de distintos produtos de acordo com as características que são disponibilizadas de forma abrangente e detalhada; a partir da classificação realizada o produto poderá ser recomendado para determinado perfil de usuário (consumidor). A Amazon.com é apresentada como a organização utilizada para realizar a avaliação da taxonomia desenvolvida, onde com base nas informações fornecidas acerca de um produto é feita a classificação deste e depois recomendado a um determinado usuário.

Chatzigeorgiou et al. (2015) propõem uma taxonomia de abordagens de avaliação que possa ser empregada em pesquisas de Engenharia de *Software*. A taxonomia proposta tem como objetivo identificar e classificar as abordagens de avaliação que são empregadas na Engenharia de *Software* com base na análise de artigos que são publicados em jornais ou conferências. De acordo com esse trabalho as categorias da taxonomia foram definidas antes da inserção dos critérios; logo, essa construção foi similar ao desenvolvimento da TaCASE. Além disso, este trabalho fornece uma definição para cada critério (ou termo) que compõe a taxonomia, como é apresentado na TaCASE, onde os critérios são definidos de acordo com a sua classe.

3. Metodologia da Pesquisa

A definição e avaliação da Taxonomia de Critérios para Avaliação de Software Educativo (TaCASE) compreendeu quatro etapas de acordo com o recomendado na literatura [Campos e Gomes, 2007], a saber: Etapa 1 Captura do conhecimento, Etapa 2 Análise dos documentos e informações, Etapa 3 Elaboração da estrutura classificatória da taxonomia e Etapa 4 Validação. As Etapas 1 e 2 foram realizadas previamente e estão descritas em Brito Junior et al. (2016a) e Brito Junior et al. (2016b), respectivamente. Na Etapa 1 realizou-se uma revisão exploratória de literatura a partir da qual foram identificadas 14 abordagens para avaliação de *software* educativo¹. Tais publicações datam de 1987 a 2015 e apresentam quantitativos de critérios que variam de 9 a 68 itens, sendo estes relativos à usabilidade e à qualidade do *software*. Na etapa 2, no conjunto total de 413 critérios, extraídos das 14 abordagens contempladas no estudo, foi analisada a equivalência, a similaridade e a redundância, tendo sido mantidos no estudo 27 critérios. Estes foram mapeados de acordo com os modelos, normas e heurísticas amplamente difundidos na Engenharia de *Software* e na Engenharia de Usabilidade.

¹[Mucchielli 1987]; [Rocha et al. 1993]; MAQSE [Campos 1994]; [Reeves 1994]; TICESE [Gamez 1998]; ASE [Vieira 1999]; IAQSEM [Gladcheff 2001]; MAEP [Silva 2002]; LORI [Nesbit et al. 2003]; MAQSEI [Atayde 2003]; FASE [Webber et al. 2009]; [Savi 2010]; [Rodrigues 2014]; e PETESE [Coomans et al. 2015]. As referências completas estão disponíveis em: <https://tinyurl.com/y8lqn23x>

A Etapa 3 teve como objetivo a elaboração da estrutura classificatória da taxonomia proposta tendo como base os resultados obtidos nas etapas anteriores. Duas classes foram definidas: uma para agrupar os critérios de qualidade de *software* (12 itens) e outra para os critérios de qualidade de uso (ou usabilidade) (21 itens), onde cada conceito (critério) apresenta uma definição. Participaram desta etapa: um aluno de pós-graduação em Ciência da Computação (nível doutorado) com experiência na avaliação de *Softwares* Educativos e dois professores pesquisadores que atuam nas áreas de Qualidade de *Software* e de Uso (Usabilidade).

A Etapa 4 destinou-se à avaliação da adequação da taxonomia junto aos potenciais avaliadores de *Softwares* Educativos, a partir da aplicação de questionário *on-line*. Para cada classe da taxonomia, um questionário foi elaborado² seguindo a mesma estrutura de 3 seções, a saber: (i) termo de consentimento de participação da pesquisa (1 item); (ii) identificação do perfil do respondente (formação e atuação profissional dos respondentes, 6 itens); (iii) avaliação da adequação da taxonomia com a apresentação do termo/conceito. Para cada item da taxonomia o respondente deveria indicar concordância ou discordância, sendo possível, para o segundo caso, indicar sugestões de mudanças. Os questionários foram difundidos nas listas de *e-mail* da Sociedade Brasileira de Computação destinada à comunidade científica que estuda Informática na Educação, Engenharia de *Software* e Interação Humano-Computador e ficaram disponíveis de 28/06/2017 a 14/07/2017.

4. Taxonomia de Critérios para Avaliação de Software Educativo (TaCASE)

Para a concepção da Taxonomia de Critérios para Avaliação de *Software* Educativo, foram consideradas as nomenclaturas e definições amplamente utilizadas na Engenharia de *Software*. Para os critérios da classe de qualidade de *software* tomou-se como base a terminologia da ISO 9126-1 (2003) e do *Swebok Guide* (2003), enquanto que para os critérios de qualidade de uso (usabilidade) considerou-se o vocabulário da ISO 9241-1 e das heurísticas definidas por Nielsen (1994), Shneiderman e Plaisant (2004) e Bastien e Scapin (1993). Fez-se necessário eliminar as duplicidades existentes entre os modelos para favorecer o uso homogêneo de termos e de suas respectivas definições. Além disso, tendo em vista a importância de manter a rastreabilidade entre os critérios da TaCASE e as 14 abordagens de avaliação de *software* educativo contempladas no estudo, esta informação foi mantida na taxonomia proposta.

A partir da Figura 1 é possível identificar os 21 critérios de avaliação definidos em TaCASE para a qualidade de *Software*. Cada termo da taxonomia tem destaque em **negrito** e é seguido da definição adotada. Em adicional, entre parêntese e em *itálico*, tem-se a lista com os nomes das abordagens para avaliação de *software* educativo que contempla o respectivo critério. Quatro critérios (*Acurácia*, *Conformidade Funcional*, *Tolerância a Falhas* e *Maturidade*) não possuem correspondentes nas abordagens para avaliação de SE, mas foram mantidos diante da importância destes para a qualidade de qualquer tipo de *software*, independente de ser ou não educativo. O mesmo formato foi adotado para os 12 critérios de avaliação definidos em TaCASE para a Qualidade de Uso (Usabilidade) (Figura 2).

² Questionários para avaliação dos critérios de qualidade de *software* (<https://tinyurl.com/ybdoa4zf>) e de qualidade de uso da TaCASE (<https://tinyurl.com/ychuzejh>).

A organização da taxonomia para a classe de Qualidade de *Software* considera um subgrupo dos termos/conceitos (Figura 1), onde: na primeira coluna têm-se uma concentração dos conceitos mais presentes na fase de concepção do *software*, na segunda coluna se encontram aqueles mais próximos da fase de uso do sistema, e na terceira coluna, os critérios que são relevantes para permitir uma boa evolução do *software*.

Taxonomia de Critérios para Avaliação de Software Educativo (TaCASE)		
Critérios de Qualidade de Software		
<p>Adequação: Capacidade de prover um conjunto apropriado de funções para tarefas e objetivos dos usuários especificados (REEVES, FASE)</p> <p>Acurácia: Capacidade de prover, com o grau de precisão necessário, resultados ou efeitos corretos ou conforme acordados</p> <p>Conformidade Funcional: Capacidade de estar de acordo com normas, convenções ou regulamentações previstas em leis e prescrições similares relacionadas à funcionalidade.</p> <p>Tempo de resposta: Capacidade de fornecer tempos de resposta, de processamento e de taxas de transferência apropriados, ao executar suas funções nas condições estabelecidas (MAQSE, Rocha, Rodrigues)</p> <p>Recursos: Capacidade de usar tipos e quantidades apropriados de recursos, quando o software executa suas funções sob condições estabelecidas (MAQSE, Rocha, MAQSEI)</p> <p>Conformidade Portátil: Capacidade de estar de acordo com normas ou convenções relacionadas à portabilidade (MAQSE, LORI, Rocha, FASE)</p> <p>Interoperabilidade: Capacidade de interagir com um ou mais sistemas especificados (PETESE)</p>	<p>Operacionalidade: Capacidade de possibilitar ao usuário operá-lo e controlá-lo (MAQSE, Rocha, MAEP)</p> <p>Inteligibilidade: Capacidade de possibilitar ao usuário compreender se o software é apropriado e como ele pode ser usado para tarefas e condições de uso específico (Rocha)</p> <p>Aprensibilidade: Capacidade de possibilitar ao usuário entender como ele funciona e aprender como usar a aplicação (Reeves, LORI, Rocha, IAQSEM, PETESE)</p> <p>Tolerância à Falhas: Capacidade de evitar falhas e manter um nível de desempenho especificado em casos de defeitos no software ou de violação de sua interface especificada.</p> <p>Recuperabilidade: Capacidade de restabelecer seu nível de desempenho especificado e recuperar os dados diretamente afetados no caso de uma falha (TICESE, MAEP)</p> <p>Segurança de Acesso: Capacidade de proteger informações e dados, de forma que pessoas ou sistemas não autorizados não possam lê-los nem modificá-los e que não seja negado o acesso às pessoas ou sistemas autorizados (MAQSE, IAQSEM)</p>	<p>Estabilidade: Capacidade de evitar efeitos inesperados decorrentes de modificações no software (MAQSE, Rocha)</p> <p>Analisabilidade: Capacidade de permitir o diagnóstico de deficiências ou causas de falhas no software, ou a identificação de partes a serem modificadas (MAQSE, Rocha)</p> <p>Adaptabilidade: Capacidade de ser adaptado para diferentes ambientes especificados, sem necessidade de aplicar outras ações ou meios além dos fornecidos para tal finalidade (Reeves, TICESE, IAQSEM, ASE, Rodrigues, FASE)</p> <p>Modificabilidade: Capacidade de permitir que uma modificação especificada seja implementada (MAQSE, Rocha)</p> <p>Substituibilidade: Capacidade de ser usado em substituição a outro produto de software especificado, com o mesmo propósito e no mesmo ambiente (LORI, Rocha, MAEP)</p> <p>Instalabilidade: Capacidade para ser instalado em um ambiente especificado (MAQSE, IAQSEM, ASE, PETESE)</p> <p>Testabilidade: Capacidade de permitir que o software, quando modificado, seja validado (MAQSE, TICESE, PETESE)</p> <p>Maturidade: Capacidade de evitar falhas decorrentes de defeitos no software.</p>

Figura 1. TaCASE para Qualidade de Software

Os critérios *Tolerância a Falhas* (Figura 1) e *Prevenção de Erros* (Figura 2), embora pareçam semelhantes, distinguem-se por definição. O primeiro se refere ao sistema se manter em uso mesmo em caso de falhas ou defeitos, enquanto o segundo consiste na disposição de mecanismos que sejam capazes de detectar e prevenir erros causados a partir das ações dos usuários.

O critério *Adaptabilidade* (Figuras 1 e 2) diz respeito à possibilidade de mudança de ambiente onde o sistema opera (Figura 1), mas também sobre como o sistema se adapta para corresponder às necessidades e características de seus usuários (Figura 2). Considerando que as classes da taxonomia se destinam a especialistas de áreas distintas de formação, pode-se compreender que não há sobreposição de termos/conceitos neste caso.

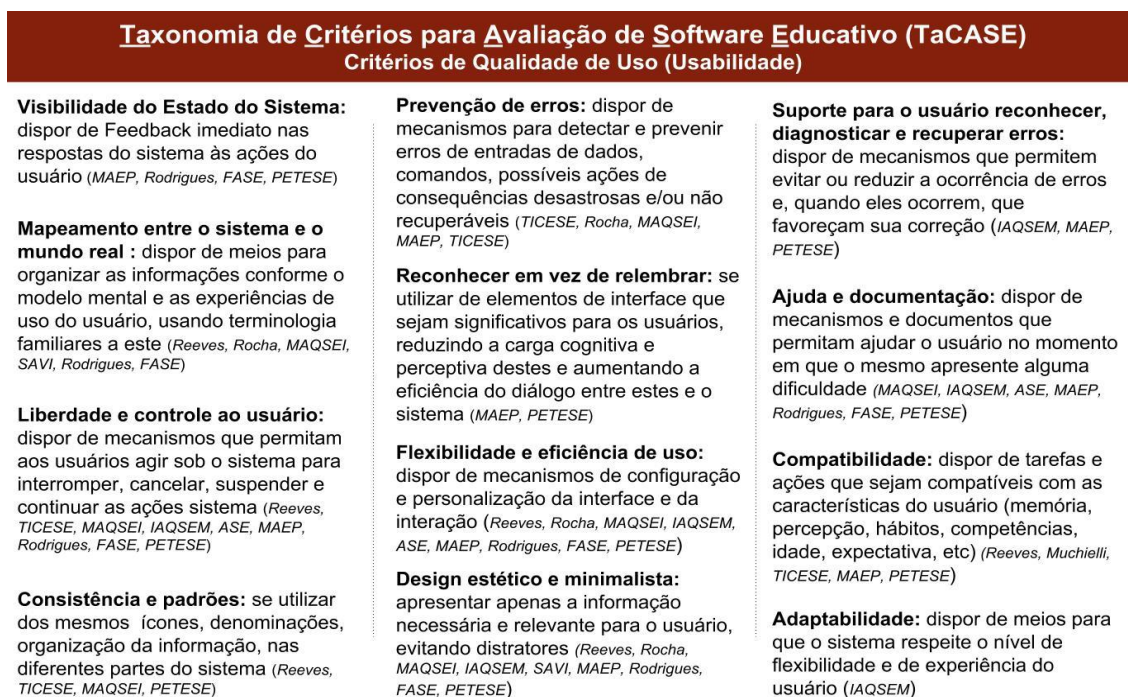


Figura 2. TaCASE para Qualidade de Uso (Usabilidade)

5. Avaliação de TaCASE por Especialistas em Engenharia de Software e Usabilidade

Para o questionário sobre Qualidade de *Software* foram obtidas 21 respostas válidas fornecidas pelos especialistas de Engenharia de *Software* e 27 pelos especialistas de Usabilidade. Ao todo foram descartados cinco questionários cujos respondentes tinham graduação incompleta ou cuja formação não era correlata à Computação.

Dos respondentes especialistas em Engenharia de *Software*, 52% possuíam doutorado, 24% mestrado e 22% graduação completa. Dos respondentes especialistas em Engenharia de *Software*, 52% possuíam doutorado, 24% mestrado e 22% graduação completa. Destes, 43% atuam no mercado de desenvolvimento de *software*, 43% se dedicam à atividade de ensino lecionando disciplinas de Engenharia de *Software*, Engenharia da Computação e Avaliação de *Software*. Os 14% restantes são pós-graduandos nos níveis de mestrado ou doutorado.

O quantitativo da **concordância** dos respondentes especialistas em Engenharia de *Software* para a relação Critério/Conceito proposta em TaCASE segue a distribuição presente no Gráfico 1. Os critérios cujos respondentes apontaram mais discordância estão com destaque em vermelho. Dentre as proposições de melhoria fornecidas pelos participantes, têm-se para: (i) *Instabilidade* → especificidade, instalabilidade, capacidade, portabilidade, interoperabilidade ou conformidade de instalação; e definição sugerida: “capacidade do *software* de ser utilizado sem falhas ou defeitos; (ii) *Maturidade* → robustez e definições sugeridas: “capacidade do *software* de não apresentar falhas” ou “capacidade do *software* de não apresentar defeitos”; (iii) *Aprensibilidade* → inteligibilidade, amigabilidade ou indutividade; e definições sugeridas: “capacidade do produto de *software* de possibilitar ao usuário entender como ele funciona e/ou aprender a usá-lo” ou “capacidade de aprendizado do produto de *software*”; (iv) *Tolerância a Falhas* → tolerância à falta, robustez ou maturidade; e

definição sugerida: “especificadas em suas definições possíveis falhas como queda da rede e falta de energia”; e (v) *Tempo de Resposta* → responsividade ou desempenho.



Gráfico 1. Qualidade de Software de TaCASE e o quantitativo de concordância sob a perspectiva do especialista em Engenharia de Software

Mesmo para os itens com maior concordância entre os respondentes, algumas sugestões de melhoria foram indicadas. Mudança de termo para os critérios: *Segurança de acesso* → segurança; *Recuperabilidade* → manutenibilidade; *Operacionalidade* → facilidade de operação do *software*, controlabilidade ou operabilidade; *Adaptabilidade* → interoperabilidade; *Conformidade Funcional* → legalidade ou conformidade legal; *Analisabilidade* → auditabilidade; *Recursos* → escalabilidade ou eficiência em relação aos recursos; e, *Conformidade portátil* → portabilidade.

No tocante às sugestões para melhoria das definições foram apontadas: *Adequação*: “capacidade do *software* de atender com qualidade às expectativas (sejam elas tarefas/objetivos) dos usuários”; *Interoperabilidade*: “capacidade de interagir com diferentes sistemas”, “capacidade de interagir com um ou mais sistemas especificados” ou “capacidade de implementar de forma válida um ou mais padrões tecnológicos especificados”; *Intelegibilidade*: “capacidade de permitir que o usuário realize suas tarefas em várias condições de uso”; *Testabilidade*: “capacidade de permitir que o *software*, quando modificado, seja testado”; *Conformidade portátil*: “capacidade de ser desenvolvido, implantado e usado em uma plataforma diferente da originalmente especificada”.

Dentre os respondentes especialistas em Engenharia de Usabilidade, 29,6% são doutores, 40,7% são mestres e 29,7% são graduados em Ciência da Computação e áreas afins. Estes atuam como desenvolvedores de *software* (44,4%), professores de nível superior (37%), especialistas em *User eXperience (UX)* e Interação Humano-Computador (IHC) (7,4%); na área de informática na educação (7,4%) ou são estudantes de pós-graduação (3,7%).

O quantitativo da **concordância** dos respondentes especialistas em Qualidade de Uso para a relação Critério/Conceito proposta em TaCASE segue a distribuição descrita no Gráfico 2. Percebe-se um alto nível de concordância para esta classe da taxonomia. A menor concordância se deu para o *Mapeamento entre o sistema e o mundo real* (21 respondentes em acordo, 77,77%) com a indicação de mudança de nomenclatura para modelo cognitivo e sugestão de modificar a descrição de forma a deixar explícito que “o docente deve ser o responsável por indicar o caminho ao aprendiz, mesmo quando se apoia no uso de um *Software* Educativo”, assim como sobre “a importância de obter opinião de especialistas de diferentes áreas com a finalidade de assegurar um produto de qualidade tanto técnica quanto pedagógica”.

Quantitativo de concordância com os critérios de qualidade de uso de TaCASE

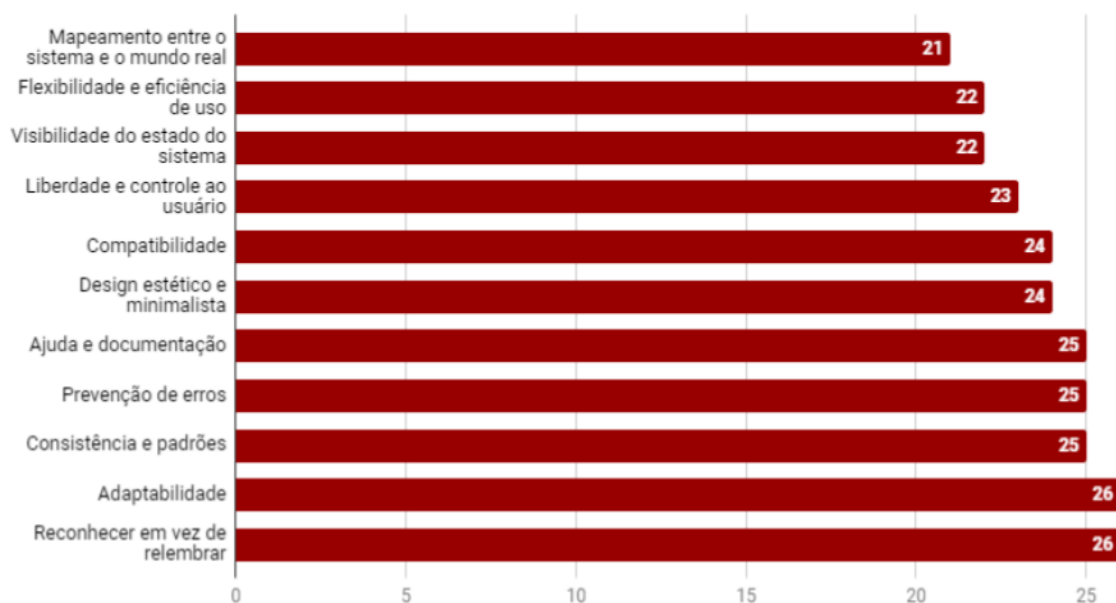


Gráfico 2. Qualidade de Uso de TaCASE e o quantitativo de concordância sob a perspectiva do especialista em Engenharia de Software

Mesmo para os itens com maior concordância entre os respondentes, algumas sugestões de melhoria foram indicadas, por exemplo: Mudança de termo para os critérios: *Prevenção de erros* → tolerância a falhas; *Design estético e minimalista* → atratividade; *Compatibilidade* → adaptabilidade; *Liberdade e controle ao usuário* → operabilidade; e, *Visibilidade do estado do sistema* → resposta do sistema.

No tocante às sugestões para melhoria das definições foram apontadas: *Reconhecer em vez de lembrar*: diz respeito a todos os elementos da interface que têm um papel importante na redução da carga cognitiva e no aumento da eficiência do diálogo.; e *Adaptabilidade*: diz respeito aos meios implementados que agreguem fluidez de infraestrutura e permitam que o sistema respeite o nível de flexibilidade e de experiência do usuário.

6. Considerações Finais

Considerando a importância da avaliação de *Softwares* Educativos e entendendo que existem desafios referentes à diversidade de abordagens disponíveis para este propósito, sendo um deles a falta de convergência entre os critérios existentes, este trabalho apresentou o processo de concessão da Taxonomia de Critérios para Avaliação de

Software Educativo (TaCASE) e avaliação desta por especialistas de Engenharia de *Software* e de Usabilidade.

A concepção de TaCASE se deu dentro de uma pesquisa realizada ao longo de dois anos e considerou a análise de 413 critérios de avaliação de *software* educativo, extraídos de 14 abordagens disponíveis na literatura. De forma a favorecer a homogeneidade sobre o conhecimento nesta área interdisciplinar, conceitos amplamente utilizados na Engenharia de *Software* e na Engenharia de Usabilidade foram adotados para a proposição de TaCASE. A taxonomia contempla duas classes que agrupam 21 critérios de qualidade de *software* e 12 critérios de qualidade de uso (usabilidade).

A fim de avaliar a taxonomia proposta, foram aplicados questionários com especialistas de Engenharia de *Software* e de Usabilidade, a fim de identificar a concordância (ou discordância) destes diante dos critérios e definições propostos. Foram obtidas 48 respostas válidas, 27 para o primeiro perfil e 21 para o segundo, com a difusão dos questionários em listas de *e-mail* da comunidade científica que ficaram disponíveis durante duas semanas. Os resultados obtidos indicam maior concordância dos participantes para os itens que compõem a classe de critérios de qualidade de uso, sendo o critério de *Mapeamento entre o sistema e o mundo real* o que apresentou menor concordância (21 respondentes de acordo, 77,77%). Para a classe de qualidade de *software*, as discordâncias foram significativas para os critérios: *Instabilidade* (11 respondentes de acordo, 52,3%); *Maturidade e Apreensibilidade* (6 respondentes de acordo, 28,5%); *Tolerância a Falhas e Tempo de Resposta* (5 respondentes de acordo, 23,8%).

Para continuação deste trabalho de pesquisa, pretende-se desenvolver um recurso digital para apresentar a taxonomia e facilitar a sua aplicação, assim como ampliá-la para incluir uma classe que considere critérios de qualidade pedagógica para SE. Além disso, será realizada uma nova análise na taxonomia com o objetivo de fazer as adaptações que foram sugeridas pelos especialistas que participaram da pesquisa e assim será feita uma nova validação. E por fim, a taxonomia será evoluída para uma ontologia.

Referências

- Bastien, C.; Scapin, D (1993). RT-0156 – Ergonomic criteria for the evaluation of human-computer interfaces. Rapport technique de l'INRIA.
- Brasil, Lina Barreto et al. Objetos de aprendizagem, competências profissionais para profissionais de saúde e e-learning: estudos para desenvolvimento de uma taxonomia. *Jornal Brasileiro de TeleSaúde*, v. 2, n. 2, p. 23-28, 2013.
- Brito Junior, O. de O., and Aguiar, Y. P. C. (2014). “Análise de abordagens objetivas para avaliação de softwares educativos”. In: Proceedings of the 13th Brazilian Symposium on Human Factors in Computing Systems. Foz do Iguaçu-PR, 2014 .
- Brito Junior, O. de O., Aguiar, Y. P. C., Tavares, T. A. (2016a). “Abordagens Objetivas para Avaliação de *Softwares* Educativos”. In: MOURA, Assis Souza de; AQUINO, Edineide Dias de. Pesquisas: teorias e práticas. João Pessoa: Penalux, 2016. v.3.
- Brito Junior, O. de O., Aguiar, Y. P. C., Tavares, T. A. (2016b). “Abordagens para Avaliação de *Softwares* Educativos e sua Coerência com os Modelos de Qualidade

- de Software”. Anais do XXVII Simpósio Brasileiro de Informática na Educação (SBIE), Uberlândia-MG, 2016.
- Chatzigeorgiou, Alexander et al. A Taxonomy of Evaluation Approaches in Software Engineering. In: Proceedings of the 7th Balkan Conference on Informatics Conference. ACM, 2015. p. 3.
- Coomans, S.; Lacerda, G. S. (2015). "PETESE, a Pedagogical Ergonomic Tool for Educational Software Evaluation". *Procedia Manufacturing*, v. 3, [p. 5881-5888], 2015.
- Da Silva, R. S. T., Silva, W. R. de A., Cardoso Filho, R. J., Pereira, W. S., Aguiar, Y. P. C., Dantas, V. F. (2016) “Aplicação comparativa de diferentes abordagens de avaliação para o software Duolingo - A complexidade de escolher uma abordagem adequada”. *Revista Tecnologias na Educação*, V. 8, N. 16.
- De França, R. S., and Da Silva, A. C. B. (2014) “Avaliação de softwares educativos para o ensino de Língua Portuguesa”. *Revista Brasileira de Informática na Educação (RBIE)*, V. 22, N. 03, P. 23.
- ISO/IEC 9214-1 (2003), Ergonomic Requirements for office work with visual display terminals – part 11.
- ISO/IEC 9126-1 (2003). Software engineering - Product quality - Part 1: Quality model. In: Associação Brasileira de Normas Técnicas, 2003.
- Marques, J., De Fátima, R. B. M., Gonçalves, P. L. (2005) “Uma proposta de taxonomia de competências para a área de computação”. In: Proceedings of XXV Congresso da Sociedade Brasileira de Computação, São Leopoldo-RS, Brasil. 2005.
- Nielsen, J (1995). “Usability engineering”. San Francisco: Morgan Kaufman.
- Peixoto, D. C. C., Campos, S. T., Resende, R. F. (2015) “Avaliação de Jogos Educacionais Multiusuários: Uma Revisão Sistemática da Literatura,” Anais do XXVI Simpósio Brasileiro de Informática na Educação (SBIE), Maceió, Brasil.
- Pereira, W. S., Cardoso Filho, R. J., Silva, W. R. D. A., et al. (2016a). “Validação de uma abordagem combinada para avaliação de Software Educativo: avanços e desafios”. *Revista Tecnologias na Educação*, V. 8, N. 16.
- Pereira, W. S., da Silva, R. S. T., Cardoso Filho, R. J., et al. (2016b) “Avaliação de Software Educativo: Análise de Abordagens para Definição de Diretrizes.”. *Nuevas Ideas en Informática Educativa*. Santiago do Chile, p. 557-562.
- Shneiderman, B. and Plaisant, C (2004). *Designing the user interface: strategies for effective human-computer interaction*. 4. ed. Addison-Wesley Publishing Company.
- Swebok (2004). "Guide to the Software Engineering Body of Knowledge". Version Project of the IEEE Computer Society Professional Practices Committee.
- Ziegler, Cai-Nicolas; Lausen, Georg; Schmidt-Thieme, Lars. Taxonomy-driven computation of product recommendations. In: Proceedings of the thirteenth ACM international conference on Information and knowledge management. ACM, 2004. p. 406-415.