

Avaliação de Software Educativo: a complexidade de escolher uma abordagem adequada

Raphael Salviano da Silva, Williane Rodrigues Silva, Raimundo Cardoso Filho, Wendell Pereira, Yuska Aguiar, Vanessa Dantas

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

{raphael.salviano, williane.rodrigues, raimundo.filho,
wendell.pereira, yuska, vanessa}@dcx.ufpb.br

Abstract. *The plurality of approaches to evaluate Educational Software (ES) requires the decision-making about which approach should be used to evaluate the ES chosen to be adopted. These approaches have criteria to verify the compliance of ES according to specific measurement tools with numerical or subjective scales, or percentage values. Considering the differences, the educator has the complex task of choosing an approach whose evaluation criteria and evaluation instruments are clear, embracing and unambiguous, and whose results are easy to analyze. The present work evaluated the Duolingo using different approaches in order to identify the complexity of each application, their scope coverage, and the quality of the results.*

Resumo. *A pluralidade de abordagens para avaliação de Software Educativo (SE) exige a tomada de decisão sobre qual abordagem utilizar na avaliação do SE que se pretende adotar. As abordagens possuem critérios para verificação da conformidade do SE a partir de escalas numéricas, subjetivas ou de valores percentuais. Diante das diferenças o educador se enfrenta a complexidade da escolha de uma abordagem cujos critérios e instrumentos de avaliação sejam claros, abrangentes e não ambíguos, e os resultados sejam fáceis de analisar. Neste trabalho descreve-se a avaliação do Duolingo a partir de diferentes abordagens, para identificar a complexidade, a abrangência e a qualidade dos resultados da aplicação de cada abordagem.*

1. Introdução

É cada vez mais comum o uso de recursos digitais no processo de ensino-aprendizagem, sendo crescente a influência da Era da Informação na educação [Costa *et al.* 2003]. Segundo Frescki (2008), a função da educação é auxiliar as pessoas na produção da sua própria realidade material e de sua consciência sobre ela. Desta maneira, os *Softwares Educativos* (SE) podem auxiliar no processo de ensino-aprendizagem a partir da contextualização de determinados conteúdos, diante da realidade de cada aluno.

O uso dos SEs tem, em potencial, uma boa aceitação, pois estes acrescentam elementos lúdicos ao processo de ensino-aprendizagem. No entanto, dificuldades podem ser enfrentadas pelos educadores na implantação de SEs no ambiente escolar. A escolha

por um SE que corresponde às expectativas do professor e às necessidades pedagógicas do conteúdo abordado exige atenção, sendo necessário garantir que o SE atenda a um conjunto de requisitos para que seu uso seja efetivo.

O SE precisa funcionar corretamente, seus elementos pedagógicos precisam estar de acordo com os propósitos dos conteúdos abordados e a interface deve ser condizente com as características, necessidades e limitações dos usuários. Problemas em qualquer um destes níveis pode prejudicar o processo de construção do conhecimento. Portanto, é necessário, antes de decidir adotar um SE, certificar-se de que este será um elemento agregador no aprendizado, avaliando-o antes de seu uso.

Para Brito e Cavalcante (2003) avaliar um SE consiste em identificar como este pode ser utilizado no âmbito educacional de modo que ajude o aprendiz a construir seu conhecimento. A diversidade de abordagens para avaliação de SE disponíveis na literatura apresenta desafios para o educador. É difícil escolher a abordagem a adotar, considerando a natureza do SE (*desktop, mobile, web*) e o *background* e *expertise* do avaliador; é possível, ainda, que a avaliação do mesmo SE por instrumentos diferentes resulte em diagnósticos distintos. Nesse contexto, este artigo objetiva identificar a influência da escolha da abordagem de avaliação de SE no diagnóstico resultante da sua aplicação, e apresentar as dificuldades encontradas durante a realização das avaliações.

O artigo está organizado em cinco seções, incluindo esta. Na seção 2 apresenta-se as abordagens consideradas no estudo. Na seção 3 descreve-se a metodologia adotada na execução das avaliações. A seção 4 contém o resultado obtidos para a avaliação do Duolingo. Por fim, na seção 5 estão as considerações finais e trabalhos futuros.

2. Métodos de Avaliação de Softwares Educacionais

A avaliação de SE deve considerar características de qualidade didático-pedagógica para que os objetivos da ergonomia de *software* e dos educadores possam convergir [Silva 1998] *apud* [Andres e Cybis 2000]. Vários autores [Rocha e Campos 1992]; [Silva e Vargas 1999]; [Oliveira *et al.* 2001] argumentam sobre a necessidade de considerar paradigmas pedagógicos na avaliação de SEs para orientar os professores na aquisição e no emprego desses recursos. Além disso, os SEs devem ser funcionalmente corretos, de acordo com a Engenharia de *Software* [Sommerville 2007], sendo importante que a interatividade e a interface oferecidas sejam adequadas ao público ao qual se destina [Preece *et al.* 2013].

Segundo Chacon *et al.* (2012) as principais abordagens objetivas (uso de *checklist*) para a avaliação de SEs são: (i) Método de Reeves [Campos 1989]; (ii) Técnica de Mucchielli [Silva 1998]; (iii) *Learning Object Review Instrument* (LORI) [Nesbit *et al.* 2002]; (iv) Técnica de Inspeção de Conformidade Ergonômica de Software Educacional (TICESE) [Gamez 1998].

A variedade de abordagens disponíveis para avaliar SEs reflete em diferenças relativas (a) ao número de critérios considerados; (b) à natureza dos critérios (pedagógicos, qualidade de *software* e de uso); (c) à natureza e granularidade das escalas adotadas (instrumentos de mensuração) e (d) à qualidade dos resultados alcançados. Logo, o educador pode se questionar sobre qual abordagem utilizar para avaliar um SE. Para familiarizar o educador sobre as abordagens citadas, segue uma breve explanação.

2.1. Método de Reeves

Consiste em 24 critérios para avaliação do SE, sendo 10 relacionados à interface (usabilidade) e 14 aos aspectos pedagógicos. Cada critério é associado a uma escala bidirecional não numérica, que se utiliza de conceitos antagônicos. Para os critérios de usabilidade, os conceitos positivos estão à direita e os negativos à esquerda. Cabe ao avaliador verificar, para cada critério, como o SE se apresenta, e realizar uma marca sobre a escala. Neste caso, havendo predominância das marcações à esquerda, percebe-se uma fragilidade do SE em relação aos aspectos de usabilidade. Para os critérios pedagógicos, os conceitos apresentados nas setas não são, necessariamente, positivos e negativos e a marcação à direita ou à esquerda não indica qualidade do SE para estes critérios. O resultado da avaliação é o conjunto de critérios e suas setas marcadas. O diagnóstico é realizado pelo educador mediante a análise da disposição dos pontos marcados nas setas.

2.2. Técnica de Mucchielli

Permite a avaliação global de *software* considerando o público alvo do SE. Propõe 10 critérios associados à escala numérica com pontuação variando entre 1 (um) e 5 (cinco). O resultado de sua aplicação é a análise da associação dos valores aos critérios, cuja apresentação é um conjunto de barras com os respectivos critérios e valores atribuídos.

2.3. Avaliação de LORI

Consiste em um *checklist* com 9 critérios (1 pedagógico, 4 de interface e 4 de qualidade de *software*). Os critérios estão associados a escalas numéricas de 5 pontos com variação de 1 (um) à 5 (cinco). O resultado obtido é semelhante ao de Mucchielli.

2.4. Técnica de TICESE

Possui três módulos: (i) classificação: formulário com informações técnicas do SE; (ii) avaliação: da documentação (com 2 critérios) e do SE (com 10 critérios); e (iii) avaliação contextual: com apenas 1 critério. Os critérios possuem um peso que variam de 0 (zero – não se aplica) à 1,5 (um e meio – muito importante).

3. Metodologia Aplicada

Para analisar, comparativamente, as abordagens para avaliação de SE apresentadas, adotou-se a seguinte estratégia: (i) Definição da equipe de avaliação: quatro discentes do curso de Licenciatura em Ciência da Computação da Universidade Federal da Paraíba (UFPB-Campus IV), sob a supervisão de duas professoras com experiência em interação humano computador, qualidade de *software* e engenharia de *software*; (ii) Experiência da equipe: dois discentes cursaram a disciplina de Softwares Educacionais oferecida pela IES, tendo experiência prévia com avaliação de SE a partir de Reeves. Todos os docentes passaram por um período de estudo teórico sobre as abordagens de avaliação, experienciando a aplicação destas com auxílio das professoras; um dos discentes é professor de língua estrangeira; (iii) Escolha do SE: o Duolingo (<https://www.duolingo.com/>) é um aplicativo para o estudo de idiomas disponível nas versões *web* e *mobile*. Seu conteúdo é agrupado em níveis (iniciante, intermediário e avançado), com temas de estudo (saudações, comidas, animais, família, etc) organizados

em lições e atividades. O nível de dificuldade das lições aumenta à medida que o aluno progride entre os níveis. Foram avaliadas as dez primeiras lições disponíveis no Duolingo, em ambas as versões a fim de verificar se as abordagens são válidas independente da plataforma do SE; (iv) Definição da dinâmica de avaliação: os alunos formaram duplas com diferentes configurações, alternando entre os membros, as versões do SE (*web* e *mobile*), e as abordagens de avaliação, de modo a evitar uma percepção tendenciosa e/ou limitada; e (v) Escolha de elementos de interesse da análise comparativa: facilidade de aplicação (tempo necessário e dúvidas frequentes), abrangência da abordagem frente aos critérios e facilidade de interpretação dos resultados.

4. Avaliação do Duolingo: versões web e mobile

Nas subseções seguintes são apresentados os resultados das avaliações realizadas. Por fim, tem-se uma análise geral e o relato das principais dificuldades enfrentadas.

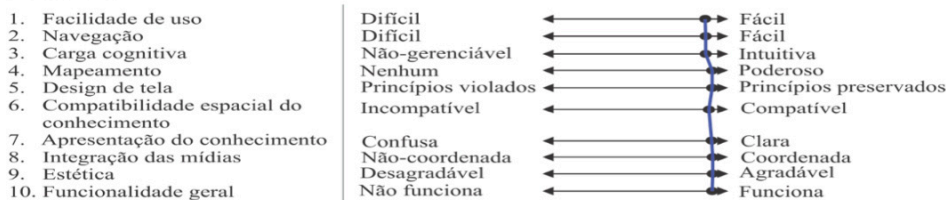
4.1. Resultado da aplicação de Reeves

Os resultados apresentam semelhanças entre as versões *web* e *mobile* (Figura 1) para todos os critérios. Em ambas as versões, os aspectos de interface foram associados a conceitos positivos. A diferença nos critérios pedagógicos (Figura 2) acontece para a acomodação das diferenças individuais (mais próximo de “*multifacetada*” para a versão *mobile*) e aprendizado cooperativo (mais próximo de “*integral*” para a versão *web*). Pode-se afirmar que os elementos pedagógicos das versões mantêm-se consistentes, possibilitando que o aluno possa alternar entre as versões sem prejuízo.

O fato de Reeves se apoiar em setas gráficas com marcação não numérica torna a análise do resultado subjetivo e difícil para avaliadores inexperientes. Sua aplicação demandou 4 horas para os avaliadores chegarem a um consenso sobre as marcações, evidenciando a essência subjetiva do método. O diagnóstico pode ser inconclusivo para os critérios pedagógicos, pois os conceitos antagônicos não são bons ou ruins.

Critérios da interface com o usuário

Versão Web



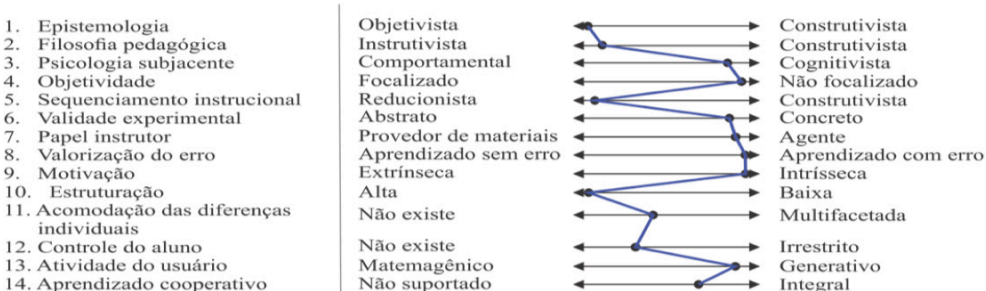
Versão Mobile



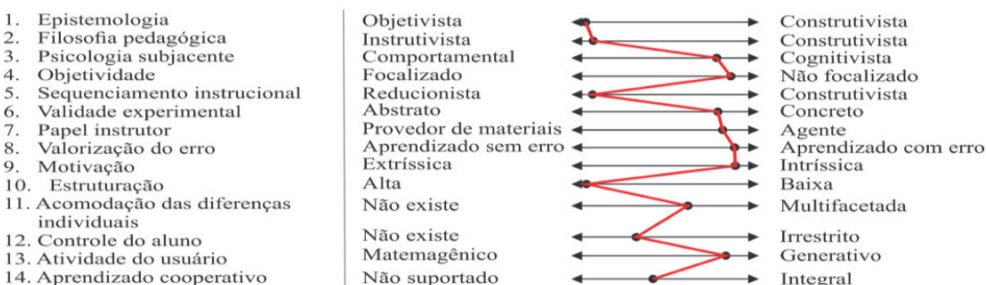
Figura 1. Aplicação de Reeves para avaliação do Duolingo nas versões Web e

Critérios pedagógicos

Versão Web



Versão Mobile



MobileFigura 2. Aplicação de Reeves para avaliação do Duolingo nas versões Web e Mobile

4.2. Resultado da aplicação de Mucchielli

Os 10 critérios de Mucchielli foram avaliados de forma positiva para ambas as versões (Figura 3) recebendo conceito máximo em sua maioria. Entretanto, dois critérios apresentam discrepâncias que influenciam na percepção do resultado: “a qualidade das ajudas” e a “qualidade do documento de acompanhamento”. Na versão web, o site possui informações úteis ao usuário, como detalhes sobre como funcionam as lições, enquanto que na mobile tem-se apenas informações técnicas (desenvolvedor, versão, etc.) disponíveis “loja” a partir da qual se adquire o aplicativo. A versão web obteve 47 pontos e a mobile 45, de um total de 50 possíveis.

Esta técnica não agrupa os critérios em pedagógicos e de usabilidade, exigindo do avaliador mais atenção ao analisar os resultados antes de afirmar que os problemas encontrados no SE são prioritariamente de interface ou pedagógicos. É interessante observar, ainda, que 7 dos 10 critérios utilizam o termo *qualidade*, que por si só é subjetivo. A qualidade percebida por avaliador experiente e principiante certamente será diferente, visto que o primeiro possui *background* de experiências anteriores.

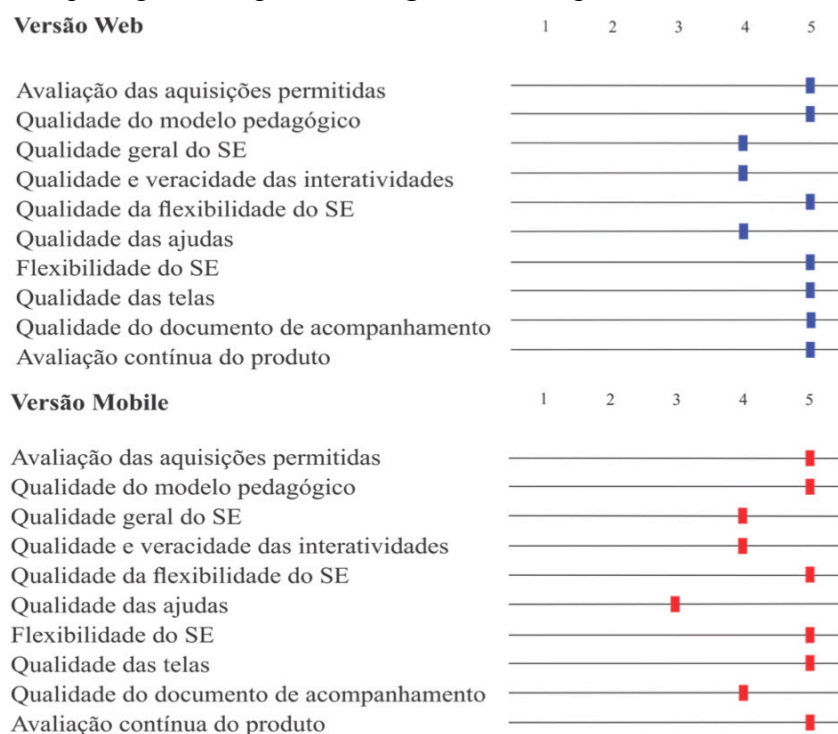


Figura 3. Aplicação de Mucchielli para avaliação do Duolingo nas versões Web e Mobile

4.3. Resultado da aplicação de LORI

O resultado diverge levemente entre as versões *web* e *mobile* (Figura 4) em relação aos critérios “*motivação*” e “*projeto e apresentação*”. Na versão *mobile*, a *motivação* é menos favorável, pois o usuário perde pontos ao cometer erros. Entretanto, nesta mesma versão, o recurso de voz utilizado para submeter respostas contribuiu para uma melhor pontuação no aspecto “*projeto e apresentação*”. Embora o somatório obtido para ambas as versões seja o mesmo (39 pontos de um total de 50), percebe-se que a versão *web* teve mais critérios com pontuação máxima. Sua aplicação demandou 3 horas.

A aplicação do método demandou um total de 2 horas para que todos os aspectos fossem checados e a equipe de avaliação chegasse a um consenso sobre a pontuação associada a cada critério - evidenciando a essência subjetiva da avaliação do SE. Além disto, percebe-se que a análise do resultado final é conclusiva, assim esta metodologia pode indicar se o software é recomendável ou não para uso.

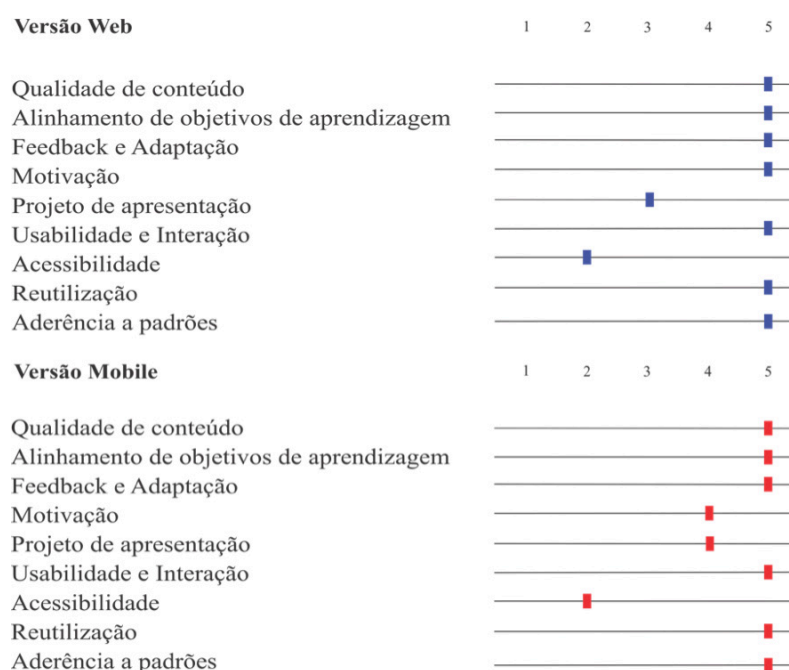


Figura 4. Aplicação de Lori para avaliação do Duolingo nas versões Web e Mobile

4.4. Resultado da aplicação de TICESE

Nas versões *web* e *mobile* os valores relativos à avaliação contextual foram idênticos. Porém algumas situações de inadequação (ou insuficiência) foram observadas. Para a versão *web* os critérios de “*dados de identificação*” (29%), de “*controle explícito*” (36,5%), de “*adaptabilidade*” (46,5%) e de “*gestão de erros*” (66%) obtiveram resultados inferiores a 75%. Na versão *mobile*, os critérios de “*dados de identificação*” (51,7%), “*adaptabilidade*” (59,5%) e “*gestão de erros*” (58%) se mantiveram como deficientes, embora os dois primeiros tenham sido melhor avaliados em relação à versão *web*. Percebe-se uma queda no critério de *avaliação da aprendizagem* (72%). Comparando com a versão *web*, percebe-se uma elevação no critério da *documentação* (70,45%) e uma diminuição para o *produto* (82,5%). Considerando a percepção da qualidade do produto pelo usuário, a versão *web* possivelmente seria considerada mais completa e melhor recomendada em relação à *mobile*.

É importante ressaltar que os valores percentuais obtidos são calculados a partir de subcritérios e de suas respectivas escalas (Figura 5). O detalhamento de critérios em subcritérios guia o avaliador e minimiza o nível de abstração para aplicar o método. No entanto, esta técnica exigiu o maior tempo de aplicação (14 horas) e o uso de planilhas auxiliares para guiar a avaliação de cada critério e realizar os cálculos exigidos pela técnica. Os resultados serem apresentados em valores percentuais pode auxiliar na interpretação do diagnóstico, pois é usual trabalhar esta forma de representação numérica.

Avaliação da Documentação

Versão Mobile

(51,75%) Dados de identificação
 (89,16%) Qualidade da informação impressa

70,45%

Avaliação do Produto

(83%) Condução
 (59,50%) Adaptabilidade
 (85%) Controle explícito
 (78%) Recursos de apoio à compreensão dos conteúdos
 (58%) Gestão de erros
 (72%) Avaliação da aprendizagem
 (92,50%) Carga de trabalho
 (82%) Significado dos códigos e denominações
 (85%) Homogeneidade
 (100%) Compatibilidade

82,50%

Avaliação Contextual

(92%) Adequabilidade

92%

Avaliação da Documentação

Versão Web

(29%) Dados de identificação
 (100%) Qualidade da informação impressa

64,50%

Avaliação do Produto

(75,25%) Condução
 (46,50%) Adaptabilidade
 (36,50%) Controle explícito
 (84%) Recursos de apoio à compreensão dos conteúdos
 (66%) Gestão de erros
 (94%) Avaliação da aprendizagem
 (93%) Carga de trabalho
 (92%) Significado dos códigos e denominações
 (93%) Homogeneidade
 (100%) Compatibilidade

88%

Avaliação Contextual

(92%) Adequabilidade

92%

Figura 5. Aplicação de Ticese para avaliação do Duolingo nas versões Web e Mobile**4.7. Análise geral dos resultados**

De modo geral, as avaliações foram positivas independente da plataforma do SE e da abordagem aplicada. Entretanto, alguns aspectos qualitativos, muitas vezes implícitos nos diagnósticos resultantes, merecem destaque. Na Tabela 1 tem-se informações relevantes para esta pesquisa considerando as (i) dimensões consideradas pelos critérios adotados nas abordagens (pedagógico, de usabilidade e de qualidade de software); (ii) a dificuldade de compreensão dos critérios, suas escalas e interpretação do diagnóstico, e (iii) do tempo necessário para aplicação das abordagens.

Tabela 1. Análise comparativa entre as abordagens

		Abordagens aplicadas para avaliação do Duolingo (Web e Mobile)			
		Reeves	TICESE	Mucchielli	LORI
Abrangência dos critérios	Pedagógico	14	3	3	1
	Usabilidade	10	3	7	4
	Qualidade de SW	0	7	0	4
Quantidade Total de critérios		24	13	10	9

Dificuldade para compreender	Critérios	Média	Média	Média	Média
	Instrumentos de medição	Baixa	Alta	Baixa	Baixa
	Diagnóstico	Média	Média	Baixa	Baixa

Tempo empregado para	Aprendizado	2h	4h	1h	1h
	Aplicação	4h	10h	2h	2h
Tempo Total		6h	14h	3h	3h

A quantidade de critérios varia de 9 (LORI) à 24 (Reeves), o que pode representar uma carga de trabalho bastante diferenciada para os avaliadores. A escolha por uma abordagem com menos critérios pode induzir o pensamento sobre a rapidez em aplicar a avaliação. No entanto, o nível de abstração necessário para avaliação a partir de poucos critérios pode implicar no emprego de mais tempo em sua realização. O fato de algumas técnicas não considerarem critérios de qualidade do *software* (Reeves e Mucchielli) pode induzir o avaliador a escolher um SE que seja muito bom pedagogicamente, mas que apresente erros de funcionamento, o que interfere no processo ensino-aprendizagem. O mesmo equívoco poderia acontecer caso fossem adotadas técnicas que privilegiam aspectos técnicos e que dão pouca ênfase às questões pedagógicas (LORI). O avaliador poderia escolher um excelente SE, mas que deixasse a desejar como instrumento de apoio à educação. É válido ressaltar que a aplicação de cada abordagem permitiu identificar limitações em diferentes aspectos do SE. Enquanto Mucchielli identificou restrições sobre ajuda para o usuário, LORI indica deficiências na motivação e no projeto de apresentação e TICESE demonstrou ressalvas sobre o controle do usuário e à documentação do SE. Portanto, a escolha da abordagem de avaliação influencia na percepção que se tem sobre o SE, explicitando certos aspectos ou ocultando fragilidades.

Outro aspecto que não está explícito nos diagnósticos obtidos é a influência da experiência e conhecimento prévios dos avaliadores na escolha e aplicação das abordagens. A equipe envolvida na avaliação do Duolingo era composta por quatro alunos com conhecimento intermediário em avaliação de SE, e mesmo assim o tempo demandado para o estudo teórico das abordagens exigiu dedicação. A cada nova abordagem aplicada, foi necessário estudar e entender os critérios, familiarizar-se com gráficos e escalas, e ao mesmo tempo explorar o SE em busca da verificação da conformidade (ou não) aos critérios. Num contexto real, onde o avaliador dispõe de prazos para avaliar o SE antes de adotá-lo, é pouco provável a dedicação para compreender diferentes abordagens. Logo, existe uma predisposição ao uso de abordagens já conhecida ou de aplicação simples e rápida. Porém, a análise isolada dos resultados de uma única abordagem pode resultar em equívocos de julgamento.

Para o Duolingo (*mobile* e *web*), mesmo diante das diferenças presentes na Tabela 1, as quatro abordagens indicam um diagnóstico favorável para sua adoção, o que pode ser consequência deste SE ser estável e largamente utilizado para o estudo de línguas, realidade diferente para boa parte de Softwares Educativos.

5. Conclusão

A avaliação de um SE é uma etapa de extrema importância para justificar sua adoção como recurso educacional. Diversas abordagens de avaliação estão disponíveis como apoio desta atividade, mas a complexidade de seus critérios e a apresentação do diagnóstico resultantes podem interferir na decisão sobre adoção ou não de um SE.

Este trabalho relata a aplicação de quatro abordagens de avaliação para um SE de aprendizado de idiomas. A análise realizada evidencia que o SE obteve avaliações favoráveis, independente da abordagem aplicada, mas também a subjetividade dos critérios adotados nas abordagens. Este cenário é desafiador para aplicação das abordagens e interpretação dos resultados.

Destaca-se que as abordagens possuem enfoques diferentes para avaliar aspectos específicos do SE. Logo, é correto afirmar que a escolha da abordagem pode variar de acordo com o objetivo do avaliador. A escolha por uma abordagem pode resultar na ausência da observação de critérios definidos como importantes pelo avaliador, uma vez que deve haver uma comunhão entre o que se pretende avaliar e os critérios presentes na abordagem.

Considerando a experiência vivenciada, pretende-se elaborar de um guia para avaliação de SE que combine diferentes abordagens. Objetiva-se fornecer aos avaliadores um instrumento que diminua a necessidade de *expertise* e *background* dos avaliadores, diminua a abstração e subjetividade dos critérios analisados e resulte em um diagnóstico cuja interpretação seja mais clara, direta e objetiva.

Referências

- Andres, D. P. and Cybis, W. A. (2000) “Um Estudo Teórico sobre as Técnicas de Avaliação de Software Educacional”. In VI Congresso Argentino de Ciencias de la Computación.
- Brito, C. L., Almeida, I. A., and Cavalcanti, L. B. (2003). “O que se Avalia e o que é Preciso Avaliar em um Software Educativo?”. Anais do Workshop de Informática na Escola. Vol. 1. No. 1. p. 334-344.
- Campos, G. H. B. (1989). “Construção e validação de ficha de avaliação de produtos educacionais para microcomputadores”. Dissertação de Mestrado – Faculdade de Educação, Universidade Federal do Rio de Janeiro. p.110.
- Chacon E. P., Borges M. N., Silva C. R. C., Clua E. W. G. 2012. Check-List: um formulário para avaliação de softwares educativos. In *III Encontro Nacional de Ensino de Ciências da Saúde e do Meio Ambiente (ENECIÊNCIAS)*, 16 a 19 de Maio, 2012, Niterói - RJ. http://www2.ic.uff.br/~ccaetano/artigos/Artigo_T-210_ENECIENCIAS.pdf. Acessado em maio de 2015.

- Costa, V. M., Rapkiewicz, C. E., de Queirós Filho, M. G., and Canela, M. C. (2003). Avaliação de sites educacionais de Química e Física: um estudo comparativo. In Anais do Workshop de Informática na Escola. Vol. 1, No. 1, p. 545-554.
- Frescki, F. B. (2008) “Avaliação da qualidade de *softwares* educacionais para o ensino de álgebra”. Centro de Ciências Exatas e Tecnológicas, Universidade Estadual do Oeste do Paraná, campus Cascavel.
- Gamez, L. TICESE. (1998). “Técnica de inspeção de conformidade ergonômica de software educacional”. Dissertação de Mestrado submetida à Universidade do Minho. Portugal.
- Nesbit, J; Belfer, K; Leacock, T. (2002). “Learning Object Review Instrument (LORI)”, <http://cenlinu1.centennialcollege.ca/aahs/LORI/help.php>.
- Oliveira B Junior, 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*. Sociedade Brasileira de Computação.
- Oliveira, C. C. O., Costa, J. W. and Moreira, M. (2001) “Ambientes informatizados de aprendizagem: produção e avaliação de software educativo”. Papirus.
- Preece, J.; Rogers, Y. and Sharp, E. (2013). Design de Interação, além da interação humano-computador, Editora Bookman, 3ª Edição.
- Rocha, A. R. and Campos, G.B. (1992). “Avaliação de Qualidade de Software Educacional”. São Paulo, Em Aberto, n. 57, v.12.
- Silva, C. R. (1998). “Bases Pedagógicas e Ergonômicas para Concepção e Avaliação de Produtos Educacionais Informatizados”. Dissertação de Mestrado submetida à Universidade Federal de Santa Catarina. Florianópolis.
- Silva, C. R. and Vargas, C. L. S. (1999). “Avaliação de Qualidade de Software Educacional”. Anais do XIX Encontro Nacional de Engenharia de Produção e V International Congress of Industrial Engineering, Rio de Janeiro, 1999.
- Sommerville, Ian. (2007) Engenharia de Software. Ed Person Education. 8ª Edição