

Avaliação de Aplicativos Educacionais Infantis com o Método MoLEva

Alexandre C. D. Batista1, Letícia I. C. Melo1, Lucas E. de Lima Vascon1, Fernanda M. P. Freire1, Flávia L. Arantes1, André C. da Silva1,2

¹Núcleo de Informática Aplicada à Educação (NIED) – Universidade Estadual de Campinas (UNICAMP)

Rua Seis de Agosto, 50 – Cidade Universitária – Campinas – SP – Brazil

2Instituto Federal de São Paulo (IFSP) — Campus Hortolândia Av. Thereza Ana Cecon Breda, s/n, Vila São Pedro — Hortolândia — SP — Brazil

Abstract. This article aims to analyze the feasibility of applying the MoLEva method to evaluate children's games on mobile devices. The games were chosen in the PlayKids app and analyzed by three reviewers, with different levels of interface assessment experience. The results show that MoLEva can be used to evaluate children's educational applications in a satisfactory way, but with some caution related to (i) the level of experience of the evaluator, (ii) the weight assigned to the categories for the calculation of the final average and (iii) a possible superficiality in questions considered as little objective in the context of children's games.

Resumo. Este artigo busca analisar a viabilidade em aplicar o método MoLEva para avaliar jogos infantis em dispositivos móveis. Os jogos foram escolhidos no aplicativo PlayKids e analisados por três avaliadores, com diferentes níveis de experiência em avaliação de interfaces. Os resultados mostram que o método MoLEva pode ser usado para avaliação de aplicativos educacionais infantis de maneira satisfatória, mas com algumas ressalvas que estão relacionadas (i) ao nível de experiência do avaliador, (ii) ao peso atribuído às categorias para o cálculo da média final e (iii) a uma possível superficialidade em questões tidas como pouco objetivas para o contexto de jogos infantis.

1. Introdução

Informações divulgadas pelo *Kids Online* no ano de 2017 [CGI.br 2018], mostram que 11% das crianças e adolescentes realizam o primeiro acesso à internet com idade inferior a 6 anos. Ainda segundo o mesmo levantamento, 93% dos acessos se dão por *smartphones*. Esse acesso, cada vez mais precoce, gera um debate entre pesquisadores sobre os malefícios e benefícios que o uso da Internet e de aplicativos pode causar.

Balbani e Krawczyk (2011) relatam que crianças pequenas são mais impactadas pela radiação emitida pelos celulares, uma vez que a espessura da pele e do crânio

dessas é 70% menor quando comparada a de adultos. Silva e Silva (2017) ainda reforçam que o uso excessivo de tais dispositivos pode trazer consequências negativas ao desenvolvimento cognitivo das crianças, levando ao isolamento social e ocasionando crises de ansiedade. Por outro lado, sabe-se que o uso de tecnologias móveis também pode ser benéfico em relação ao aprendizado, como mostra o guia da UNESCO de 2013 [UNESCO, 2013], que incentiva os governos nacionais a utilizarem celulares como recurso educacional em sala de aula, listando 13 motivos, como por exemplo a otimização do tempo em sala de aula. Além disso, segundo Almeida (1974) apud Lopes *et al.* (2015), os jogos favorecem a habilidade cognitiva, criatividade, disciplina e diversos outros fatores associados ao desenvolvimento intelectual.

À vista da crescente discussão acerca do tema, existe um interesse em promover uma melhor análise da qualidade dos aplicativos e jogos de tecnologia móvel, buscando caracterizar o conteúdo e classificar a qualidade destes, tanto de forma pedagógica como de usabilidade.

Soad (2017) desenvolveu um método de avaliação de aplicativos educacionais denominado MoLEva (*Mobile Learning Evaluation*), o qual foi elaborado através de modelos de qualidade, parâmetros métricos e critérios de julgamento, tendo principalmente como base a norma ISO/IEC 25000. Trata-se de um método para avaliar especificamente a qualidade de aplicativos educacionais móveis (*m-learning*). Para validar o MoLEva, foram realizados dois estudos de caso; o primeiro consistiu na aplicação do método para avaliar o aplicativo do ENEM (Exame Nacional do Ensino Médio); o segundo consistiu na avaliação de aplicativos para o ensino de idiomas – Doulingo e Wlingua.

Neste artigo, apresentamos os resultados da aplicação do método MoLEva para analisar dois jogos do aplicativo infantil *PlayKids*. Como contribuições, destacamos a avaliação dos jogos em si e a utilização do método MoLEva em um domínio no qual ele ainda não foi aplicado, o que nos permitiu tecer impressões a respeito do uso desse método especificamente para avaliação de jogos infantis.

2. Metodologia

O método MoLEva foi o escolhido para este estudo por tratar especificamente de aplicativos educacionais móveis. Já o *PlayKids* foi escolhido após resultados obtidos de um questionário aplicado com pais e mães de crianças de 2 a 6 anos, o qual avaliou a interação de crianças com tecnologias móveis e identificou que o aplicativo em questão é o mais utilizado pelo referido público alvo, no contexto de jogos infantis [SILVA; ARANTES; FREIRE, 2019].

Após análise do *PlayKids*, foram definidos dois jogos para a avaliação, os quais chamamos de (i) Caligrafia e (ii) Baú de Brinquedos. Antes de aplicar o método MoLEva, os jogos foram explorados pelos avaliadores, que fizeram relatos descritivos de usabilidade, conteúdo, erros e jogabilidade. Depois de se familiarizarem com os jogos, os avaliadores utilizaram o questionário final do método MoLEva para caracterizar de forma quantitativa a qualidade dos aplicativos. Os testes foram realizados por 1 avaliador experiente e 2 inexperientes, todos do meio acadêmico. Entende-se como experientes, neste estudo, aqueles que têm familiaridade com a avaliação de aplicativos e o fazem a um considerado período de tempo (3 anos em média) e inexperientes, aqueles que não realizam esse tipo de análise habitualmente. Essa abordagem foi utilizada uma vez que se busca comparar os resultados obtidos entre

avaliadores experientes (os quais espera-se que sejam mais criteriosos na avaliação) e inexperientes.

Em um encontro presencial, os avaliadores estudaram o método MoLEva, tendo como referência a dissertação de Soad (2017), pois nenhum deles havia realizado avaliações com o método anteriormente. De acordo com os avaliadores, o método MoLEva é intuitivo e de fácil aplicação. A condução dos testes ocorreu de maneira individual e em diferentes localidades, onde cada avaliador realizou os testes sem nenhuma interferência dos demais.

3. Os jogos analisados

Os jogos Caligrafia e Baú de Brinquedos foram os escolhidos para avaliação. O Baú de Brinquedos se assemelha a jogos de encaixe, amplamente desenvolvidos na versão física para o público infantil e o Caligrafia, indicado para crianças um pouco maiores que estão aprendendo a ler e a escrever, se assemelha a um tipo de tarefa escolar bastante conhecida. No jogo Caligrafia (Figura 1), é possível escolher qualquer letra cursiva do alfabeto, maiúscula ou minúscula, a qual é mostrada de forma pontilhada em uma pauta (como em um caderno de caligrafia), de modo que a criança possa utilizar o dedo para traçar a letra por cima dos pontos e ainda a reproduzir posteriormente. Porém, não há uma limitação do traço da criança – o traço é livre – e, portanto, é possível que a funcionalidade não opere da forma inicialmente esperada.



Figura 1. Capturas de tela do jogo Caligrafia.

No jogo Baú de Brinquedos (Figura 2), a criança é estimulada a realizar a arrumação dos brinquedos jogados no chão do quarto. Trata-se de um jogo que trabalha a coordenação motora e o reconhecimento de formas. Inicialmente todos os brinquedos se encontram no chão e há um baú onde deverão ser guardados. No baú, há sombras no formato dos brinquedos, o que dá uma pista sobre o local para onde o brinquedo deve ser arrastado/guardado. A criança então toca no brinquedo e arrasta até uma das sombras. Se o formato da sombra corresponder ao do brinquedo, automaticamente ele irá para dentro do baú, após uma breve animação. Caso contrário, o brinquedo faz um movimento de "tremer" e retorna para o chão, o que leva a criança a repetir o movimento. Ao preencher todas as sombras com os brinquedos correspondentes, uma animação indicando que o desafio foi completado é mostrada. O jogo então recomeça, todavia, a quantidade de brinquedos a serem guardados aumenta, e assim sucessivamente, a cada nível vencido.



Figura 2. Capturas de tela do jogo Baú de Brinquedos.

Os dois jogos buscam desenvolver a autonomia da criança, não sendo necessária a presença de um tutor. As tarefas ou missões cumpridas são recompensadas tendo em vista estimular as crianças a completar os objetivos dos jogos. Essa mecânica de jogos faz com que, em muitos casos, ocorra uma aprendizagem por assimilação.

4. O Método MoLEva

Conforme mencionado na introdução, o método MoLEva busca avaliar aplicativos educacionais móveis a partir de um modelo de qualidade, o qual é constituído por categorias e critérios. As categorias estabelecidas são dividas em três: i) Pedagógica, ii) Social e iii) Técnica. Posteriormente essas três categorias são subdividas em doze critérios de qualidade: i) Aprendizagem, ii) Conteúdo, iii) Interatividade, iv) Socioeconômica, v) Sociocultural, vi) Adequação Funcional, vii) Eficiência do desempenho, viii) Compatibilidade, ix) Usabilidade, x) Confiabilidade, xi) Segurança e xii) Portabilidade. Os critérios de i a iii fazem parte da categoria Pedagógica; iv e v constituem a categoria Social e os critérios de vi a xii compõem a categoria Técnica. A partir dessas três categorias e doze critérios, um questionário de avaliação composto por oitenta e cinco questões foi elaborado, contemplando questões que permitem parametrizar todos os critérios. O questionário é constituído por perguntas com respostas binárias e por grau de concordância, em que cada opção de resposta tem uma pontuação atribuída.

Um exemplo de pergunta do tipo respostas binárias é mostrado abaixo e a pontuação atribuída a cada alternativa é apresentada na Tabela 1.

[Adaptado de Soad (2017)] 36 - [Adequação funcional]: O aplicativo faz uso de diferentes tipos de mídia? (ex. vídeo, texto, áudio, imagens, entre outros).

() Sim () Não () Não se aplica () Avaliação Prejudicada

Tabela 1. Pontuação de questões do tipo binária

Resposta	Pontuação		
Sim	10		
Não	0		
Não se aplica	0		
Avaliação Prejudicada	0		

A seguir, é mostrado um exemplo de questão do tipo resposta por grau de concordância e a pontuação atribuída a cada alternativa se encontra na Tabela 2.

[Adaptado de Soad (2017)] 83 - [Portabilidade]: Como pode ser classificada a utilização do aplicativo em diferentes sistemas operacionais?

Ineficiente () () () () () Eficiente

() Não se aplica () Avaliação prejudicada

Tabela 2. Pontuação de questões do tipo grau de concordância

Grau de Concordância/Resposta	Pontuação
5	10,0
4	7,5
3	5,0
2	2,5
1	0,0
Não se aplica	0,0
Avaliação Prejudicada	0,0

A partir das pontuações dispostas nas Tabelas 1 e 2, Soad (2017) desenvolveu parâmetros métricos que resultaram em equações, as quais permitem a obtenção de notas, entre 0 e 100, para Critérios (Equação 1), Categorias (Equação 2) e uma nota geral para o aplicativo (Equação 3).

$$NQRC = \frac{(p_1 + p_2 + \dots + p_n)}{TPC}.10$$
 (1)

em que NQRC é o Nível de Qualidade por Critério; TPC é o Total de Perguntas consideradas por Critério; p é a pontuação atribuída para cada pergunta e n é o total de perguntas do critério avaliado.

$$NQCA = \frac{NQCR_1 + NQCR_2 + \dots + NQCR_n}{TCQ}$$
 (2)

em que NQCA é o Nível de Qualidade por Categoria; TCQ é o Total de Critérios de Qualidade relacionados à categoria avaliada e n é o total de critérios relacionados à categoria avaliada

$$NQA = \frac{NQP + NQS + NQT}{TC} \tag{3}$$

em que NQA é o Nível de Qualidade do Aplicativo; NQP é o Nível de Qualidade da categoria Pedagógica; NQS é o Nível de Qualidade da categoria Social; NQT é o Nível de Qualidade da categoria Técnica e TC é o Total de Categorias avaliadas. Ao final da avaliação, a partir do Nível de Qualidade do Aplicativo (NQA), o aplicativo é classificado por: se NQA < 50, a qualidade é de nível inferior; se $50 \le NQA \le 80$, a qualidade é de nível Superior.

5. Análise dos Jogos

5.1. Caligrafia

O método MoLEva foi aplicado pelos três avaliadores ao jogo Caligrafia e os Níveis de Qualidade por Critério para cada avaliador são mostrados nas Figura 3 e 4, em que o Avaliador 3 é o experiente e os demais são os inexperientes.

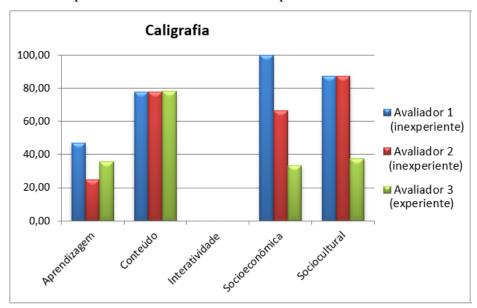


Figura 3. NQRC dos critérios de i a v para o jogo Caligrafia.

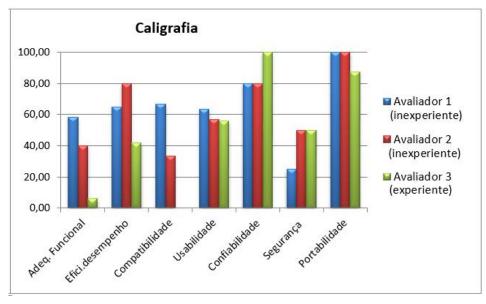


Figura 4. NQRC dos critérios de vi a xii para o jogo Caligrafia.

5.2. Baú de Brinquedos

De maneira semelhante, o método MoLEva foi aplicado pelos três avaliadores ao jogo Baú de Brinquedos e os Níveis de Qualidade por Critério para cada avaliador são mostrados nas Figura 5 e 6. A Tabela 3 resume as notas dadas pelos três avaliadores para os dois jogos.

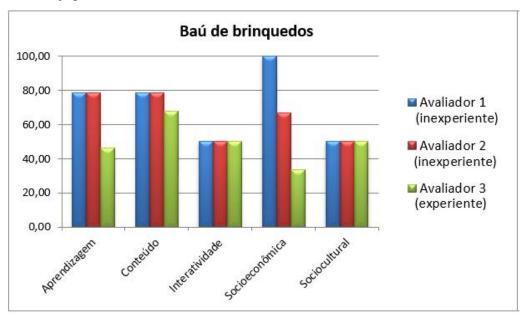


Figura 5. NQRC dos critérios de i a v para o jogo Baú de Brinquedos.

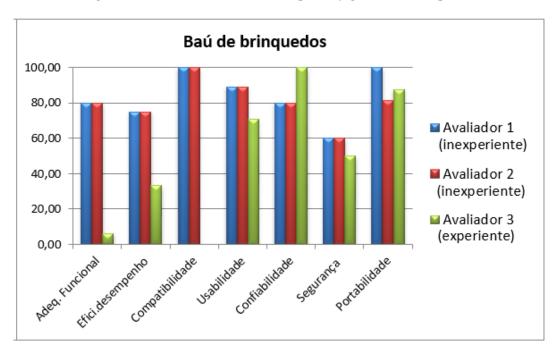


Figura 6. NQRC dos critérios de vi a xii para o jogo Baú de Brinquedos.

Tabela 3. Notas dos critérios agrupados em categorias e nota final para Caligrafia (à esquerda) e Baú de Brinquedos (à direita).

CALIGRAFIA			BAÚ DE BRINQUEDOS				
Critérios/				Critérios/			
Categorias	Avaliador 1	Avaliador 2	Avaliador 3	Categorias	Avaliador 1	Avaliador 2	Avaliador 3

Aprendizagem	46,88	25	35,71	Aprendizagem	78,57	78,57	46,43
Conteúdo	77,78	77,78	78,13	Conteúdo	78,57	78,57	67,86
Interatividade	0	0	0	Interatividade	50	50	50
Pedagógica	41,55	34,26	37,95	Pedagógica	69,05	69,05	54,76
Socioeconômica	100	66,67	33,33	Socioeconômica	100	66,67	33,33
Sociocultural	87,5	87,5	37,5	Sociocultural	50	50	50
Social	93,75	77,09	35,42	Social	75,00	58,34	41,67
Adeq. Funcional	58,33	40	6,25	Adeq. Funcional	80	80	6,25
Efici. Desempenho	65	80	41,67	Efici. Desempenho	75	75	33,33
Compatibilidade	66,67	33,33	0	Compatibilidade	100	100	0
Usabilidade	63,33	56,67	55,77	Usabilidade	89,06	89,06	70,83
Confiabilidade	80	80	100	Confiabilidade	80	80	100
Segurança	25	50	50	Segurança	60	60	50
Portabilidade	100	100	87,5	Portabilidade	100	81,25	87,50
Técnica	65,48	62,86	48,74	Técnica	83,44	80,76	49,70
NOTA FINAL	66,93	58,07	40,70	NOTA FINAL	75,83	69,38	48,71

6. Resultados e Discussão

As Figuras 3 e 4, bem como a Tabela 3, mostram que para o jogo Caligrafia a avaliação realizada pelo avaliador 3 é mais rigorosa resultando em notas menores quando comparado aos demais avaliadores. Isso é enfatizado principalmente nos critérios da categoria Técnica em que apenas a nota do critério Confiabilidade é maior do que as atribuídas pelos demais. Tal situação pode ser explicada devido a experiência e conhecimento técnico do avaliador 3, uma vez que análises contínuas de interface e conhecimento mais aprofundado fazem com que tenha um olhar mais crítico. Ainda no Caligrafia, nota-se que o critério Interatividade foi o único a receber nota zero de todos os avaliadores. Todavia, esse critério tem quatro questões e obteve como resposta, em três delas, a alternativa "Não se aplica", a qual não contribui para a nota final do aplicativo. Dessa forma, a nota geral desse critério foi constituída por apenas uma questão em que os três avaliadores assinalaram a opção "Não" que consequentemente corresponde à nota zero.

No jogo Baú de Brinquedos (Figuras 5 e 6), o avaliador 3 foi, novamente, mais rigoroso. A maior discrepância é mostrada no critério "Compatibilidade", onde os avaliadores 1 e 2 atribuíram nota 100 à funcionalidade e o avaliador 3 atribuiu nota 0. Isso pode ser explicado uma vez que a maioria das respostas para esse critério pelo terceiro avaliador tiveram sua avaliação prejudicada, sendo a única pergunta realmente respondida com um "Não", influenciando diretamente na nota zero. Outro ponto bem discrepante foi em relação à Adequação Funcional, sendo atribuídas notas iguais a 80 pelos primeiros avaliadores e 6,25 pelo terceiro. Em contrapartida, assim como na funcionalidade anterior, o avaliador 3 (especialista) atribuiu notas maiores no critério "Confiabilidade" quando comparado aos avaliadores 1 e 2.

A discrepância entre as notas atribuídas pelo terceiro avaliador pode ser explicada pela sua maior experiência como especialista em interfaces. Outra possível explicação é a interpretação dos itens avaliados - apesar de existir um exemplo para cada questão, uma interpretação incorreta ou divergente pode levar a avaliações diferentes e até mesmo comprometer a avaliação. Na Tabela 3 nota-se que apenas na

categoria Pedagógica as notas dadas pelo Avaliador 3 não são menores. Essa categoria engloba critérios mais relacionados com a forma de transmissão de conhecimento ou técnicas de aprendizagem que o aplicativo pode oferecer, o que implica uma visão mais voltada a aspectos relacionados à educação e que, neste caso, os três avaliadores compartilharam da mesma visão. Por fim, pela Tabela 3, notou-se que o avaliador mais experiente concluiu que os jogos são de nível inferior (nota final menor que 50), enquanto que os menos experientes chegaram a classificações de nível médio (nota final entre 50 e 80). Esse resultado nos mostra que a aplicação do método MoLEva para avaliação de aplicativos educacionais pode ser influenciada pela experiência e olhar crítico do avaliador, uma vez que o método não exige um número mínimo de avaliadores ou alguma experiência prévia – a avaliação feita por apenas um avaliador é considerada como válida.

Com relação ao cálculo da nota no método MoLEva, a Equação 3 mostra que todas as categorias têm peso igual na composição final da nota do aplicativo avaliado, uma vez que é realizada a média aritmética entre elas. Essa definição da nota pode ser prejudicial para a classificação final do aplicativo, já que, por exemplo, a categoria Técnica tem doze critérios enquanto a categoria Social possui apenas duas, fazendo com que um aplicativo que não tenha uma nota muito alta em Técnica, mas tem uma nota relativamente alta em Social eleve a sua média final consideravelmente. O método não foi desenvolvido para a avaliação de aplicativos infantis, mas como esses se encaixam na categoria de aplicativos educacionais móveis e pelo MoLEva ter certa abrangência, a utilização do mesmo nesse domínio pode ser considerada satisfatória, mas com algumas limitações. É importante observar, por exemplo, que, em sua primeira versão, o método foi avaliado apenas no contexto do ensino de línguas, com os aplicativos Doulingo, Wlingua e Rosetta Stone (Soad, 2017, p.89). O resultado da avaliação serviu como subsídio para o refinamento do método, o que resultou em sua versão atual. Por ter sido refinado com base nos estudos de casos que realizou, algumas questões não são aplicáveis para jogos infantis, tais como conteúdo de aprendizado, confiabilidade do conteúdo, validação de conhecimento, dentre outras. Os jogos infantis têm um modelo de ensino pouco objetivo que o MoLEva não cobre, mas ele consegue cobrir uma aplicação móvel em aspectos mais gerais (ocupação de memória, bugs, disposição dos elementos da interface etc). A utilização da resposta "Não se aplica" pode resolver o problema em um primeiro momento, mas no final pode gerar uma avaliação superficial como relatamos no caso do critério Interatividade do jogo Caligrafia, citado no início desta seção.

7. Considerações Finais

A avaliação realizada pelo método MoLEva permitiu identificar discrepâncias nas notas atribuídas por usuários experientes e inexperientes, principalmente na Categoria Técnica, o que levou a classificações de nível inferior e médio para o mesmo jogo. O conhecimento prévio do avaliador mais experiente permitiu a identificação de aspectos não tão claros para avaliadores com menos experiência. Por fim, verificamos que é possível realizar uma análise satisfatória dos jogos infantis com o método MoLEva. Apontamos ressalvas com relação (i) à experiência do avaliador, que pode influenciar a nota final do aplicativo, (ii) ao peso atribuído a cada categoria na composição da nota final e (iii) à geração de avaliações superficiais em casos em que a opção "Não se aplica" é a única cabível sucessivas vezes. Uma eventual redistribuição dos pesos das categorias e uma revisão das questões de modo a torná-las mais abrangentes reduzindo a

porcentagem de alternativas com a resposta "Não se aplica", podem tornar a avaliação mais coerente e eficaz.

Referências

- Balbani, A. P. S.; Krawczyk, A. L. (2011) "Impacto do uso do telefone celular na saúde de crianças e adolescentes". Revista Paulista de Pediatria, São Paulo, v. 29, n. 3, p. 430-436, Setembro, 2011. Disponível em: http://dx.doi.org/10.1590/S0103-05822011000300019. Acesso em: 19 fev. 2019.
- CGI.br Comitê Gestor de Internet no Brasil (2018) "TIC Kids Online Brasil Pesquisa sobre o uso da Internet por Crianças e Adolescentes no Brasil". Disponível em https://cetic.br/media/docs/publicacoes/2/tic_kids_online_2017_livro_eletronico.pdf. Acesso em: 19 fev. 2019.
- Lopes, S. R.; Paula, M.; Pólio, E.; Silva, E. N. (2015) "Jogos Enquanto Recurso do Desenvolvimento na Educação Infantil". Revista Semana Acadêmica, v.01, n.76, 2015. Disponível em: https://semanaacademica.org.br/artigo/jogos-enquanto-recurso-do-desenvolvimento-na-educacao-infantil. Acesso em 20 fev. 2019.
- Silva, L. R; Arantes, F. L.; Freire, F. M. P. (2019) "Tecnologias móveis e crianças pequenas questões iniciais". Relatório Técnico do NIED, n. 1, 2019. Campinas, Brasil. No prelo. Estará disponível em: https://www.nied.unicamp.br/biblioteca/relatorios-tecnicos/
- Silva, T. O.; Silva, L. T. G. (2017) "Os impactos sociais, cognitivos e afetivos sobre a geração de adolescentes conectados às tecnologias digitais". Revista Psicopedagogia, São Paulo, v. 34, n. 103, p. 87-97, 2017. Disponível em: http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S0103-84862017000100009&lng=pt&nrm=iso. Acesso em 19 fev. 2019.
- Soad, G. W. (2017) "Avaliação de qualidade em aplicativos educacionais móveis". 147 p. Dissertação (Mestrado em Ciências Ciências de Computação e Matemática Computacional) Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos SP. 2017.
- UNESCO (2013) "Diretrizes de políticas da UNESCO para a aprendizagem móvel". Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000227770. Acesso em 20 fev. 2019.