Ano 10 – Número/vol.31 – dezembro/2019

Revista Tecnologias na Educação

ISSN: 1984-4751

O Método MoLEva para avaliação de aplicativos educacionais infantis – um olhar considerando diferentes perfis de avaliadores

Flávia L. Arantes[1]

Alexandre C. D. Batista[1],

Letícia I. C. Melo[1],

Lucas E. de Lima Vascon[1],

Fernanda M. P. Freire[1],

André C. da Silva[1,2]

RESUMO

Neste artigo apresentamos os resultados da aplicação do método MoLEva para analisar jogos educacionais infantis em dispositivos móveis, considerando diferentes perfis de avaliadores. A avaliação nos permitiu tecer impressões e sugestões a respeito do uso do método para o contexto de jogos educacionais. Os resultados mostram que o método pode ser usado para avaliação de aplicativos educacionais apesar de estar mais voltado a aspectos técnicos. Verificamos ainda que o perfil dos avaliadores pode influenciar diretamente na avaliação devido a fatores como: experiência no universo infantil, experiência com avaliação de aplicativos e entendimento das questões.

Palavras-chave: Avaliação. Jogos. Educação.

1. Introdução

A habilidade demonstrada pelas crianças contemporâneas no manuseio das tecnologias, especialmente as móveis, causa surpresa a muitos adultos. As crianças, desde

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

[2] Instituto Federal de São Paulo (IFSP) – Campus Hortolândia

cedo, participam de diferentes práticas mediadas por tecnologias, tanto na escola quanto em casa: tiram fotos, fazem pequenos vídeos, assistem a desenhos, brincam com aplicativos. 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, Amante (2007) apresenta um conjunto de razões que justifica a integração de tecnologias ao currículo pré-escolar. Segundo a autora, as tecnologias representam um importante apoio para o desenvolvimento da linguagem, o letramento e a aquisição da escrita, o pensamento matemático, conhecimento de mundo e educação para a diversidade. Adverte, porém, que a melhoria da qualidade da educação depende, dentre outros fatores, de se saber colocar a tecnologia a serviço de um projeto educativo renovado, o que implica o engajamento de pessoas e de instituições comprometidas e preparadas. Estudos como esse estimularam a inserção do uso de TDIC (Tecnologias Digitais de Informação e Comunicação) nas diretrizes da educação infantil de alguns municípios, como é o caso de São Paulo (SMESP, 2015).

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.

Em um trabalho anterior (Batista, et al., 2019), 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 havia sido aplicado. O resultado foi pautado na avaliação dos jogos Caligrafia e Baú de Brinquedos, feita com três avaliadores: um experiente e dois inexperientes. Entende-se como experientes, 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.

Nos resultados, identificamos discrepâncias nas notas atribuídas pelo avaliador experiente quando comparadas às notas atribuídas pelos avaliadores inexperientes. Parte da discussão foi norteada pelas comparações entre as análises realizadas pelo avaliador experiente (que foi mais criterioso) e os avaliadores inexperientes, o que nos levou a concluir que a aplicação do método MoLEva para avaliação de aplicativos educacionais pode consideravelmente ser influenciada pela experiência e olhar crítico do avaliador.

Neste artigo, trazemos uma versão estendida da avaliação realizada em Batista e coletas (2019), com uma análise mais completa, que se pauta em avaliadores com perfis mais variados, tanto tecnicamente quanto com relação à experiência com crianças. Os resultados ratificam o fato do método MoLEva ser notadamente sensível ao perfil do avaliador. Além disso, as novas avaliações nos permitiram tecer impressões e sugestões mais embasadas a respeito do uso do método para o contexto de jogos educacionais infantis.

2. Embasamento teórico

2.1 A Cultura Digital e a BNCC

Há um crescente reconhecimento por parte de organismos internacionais e governos de que as TDIC são componentes fundamentais para o desenvolvimento socioeconômico e, consequentemente, para o alcance dos Objetivos do Desenvolvimento Sustentável (ODS) da Agenda 2030 da Organização das Nações Unidas (ONU). A adoção das tecnologias digitais torna-se uma variável relevante na agenda de desenvolvimento de longo prazo, sobretudo, na eliminação das possíveis fontes geradoras de contrastes e disparidades sociais (CGI.br, 2018). Não é atoa que a Cultura Digital e o Pensamento Computacional estão tão presentes na Base Nacional Comum Curricular (BNCC), como recursos transversais nas diversas áreas do conhecimento para "se comunicar, acessar e disseminar informações, produzir

conhecimentos, resolver problemas, exercer protagonismo e autoria na vida pessoal e coletiva" (Brasil, 2018, p. 9).

Neste trabalho, interessa-nos sobretudo a faixa etária entre 2 e 6 anos por ser a idade em que as crianças começam a manipular os dispositivos móveis e, consequentemente, a terem acesso aos jogos digitais educacionais. Nessa faixa etária, as crianças estão na Educação Infantil e, apesar da Cultura Digital estar longe de ser um foco nas escolas, a BNCC sugere algum contato ou manipulação por parte das crianças, para que as mesmas possam reconhecer os recursos tecnológicos como fontes de informação e audiovisuais.

> Conhecer e manipular materiais impressos e audiovisuais em diferentes portadores (livro, revista, gibi, jornal, cartaz, CD, *tablet*, etc.) (Brasil, 2018, p. 50).

Embora a BNCC esteja alinhada com os saberes da Cultura Digital, é em casa (e não na escola) que as crianças pequenas normalmente têm acesso às tecnologias, principalmente as móveis. Uma pesquisa realizada por nós, com pais de crianças entre 2 e 6 anos, mostra que todas as famílias da amostra têm celulares em suas casas, 84% têm notebook e várias famílias também possuem Smart TV e tablet (Silva, Arantes e Freire, 2019). A pesquisa também mostra que 58,3% das crianças usam tecnologias móveis para fins educacionais.

2.2. O PlayKids e os jogos analisados

O PlayKids é um aplicativo que oferece jogos, atividades e conteúdos educativos para crianças pequenas. A escolha de jogos do PlayKids para avaliação se deu 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 e Freire, 2019).

Dentro do PlayKids, os jogos Caligrafia e Baú de Brinquedos foram os escolhidos para avaliação por serem abrangentes para a faixa etária de 2 a 6 anos e por estarem disponíveis na versão gratuita.

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 Caligrafia (Figura 1), é possível escolher qualquer letra de forma ou 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 Caligrafia.

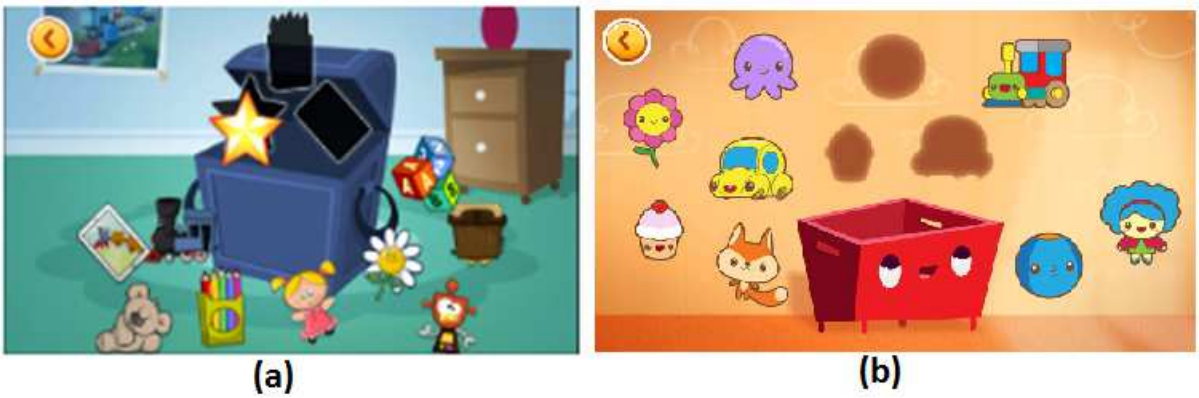
Fonte:Play kids

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.

Cabe mencionar que houve uma mudança na interface do Baú de Brinquedos ao realizar este novo estudo – verificamos que uma das versões do Playkids era 4.19.4 e outra era

4.7.4. A nova versão é mais limpa, foram removidos alguns elementos da interface (o criado mudo e o quadro), o que facilita a criança a identificar os objetos e os encaixes. A Figura 2(a) mostra o jogo na versão antiga e a Figura 2(b) mostra a versão mais nova.

Figura 2. Capturas de tela do jogo Baú de Brinquedos



Fonte:Play kids

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.

2.3. O Método MoLEva

O método MoLEva foi o escolhido para este estudo por tratar especificamente de aplicativos educacionais móveis. Conforme mencionado na introdução, o método 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

1 2 3 4 5

( ) 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).

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

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.

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

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

$$NQA = \frac{NQP + NQS + NQT}{TC} \quad (3)$$

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 ≤ NQA ≤ 80, a qualidade é de nível médio e se NQA > 80, a qualidade é de nível Superior.

3. Perfil dos Avaliadores e Metodologia

Conforme explicado na introdução, neste artigo trazemos uma versão estendida da avaliação realizada em Batista e colegas (2019). Grande parte do resultado anterior se pautou nas comparações entre as análises realizadas pelo avaliador experiente (que foi mais criterioso) e os avaliadores inexperientes. Neste artigo, apresentamos uma análise mais completa, que se pauta em avaliadores com perfis mais variados, descritos a seguir:

- Avaliador 1: atua na área de exatas (Engenharia Química), não tem experiência com crianças e nem com a avaliação/design de aplicativos.

- Avaliador 2: atua na área de exatas (Engenharia Química), tem experiência de um ano com crianças de 2 a 7 anos na educação não formal (brinquedoteca). Não tem experiência com design ou avaliação de aplicativos.
- Avaliador 3: atua na área de exatas (Engenharia Elétrica), tem experiência de 3 anos na avaliação de interfaces (avaliação heurística, percurso cognitivo e avaliação com usuários), tem experiência no design de interfaces para aplicativos móveis, mas não especificamente para jogos educacionais infantis. Não tem experiência com crianças.
- Avaliador 4: atua na área de humanas (Ciências Sociais), trabalhou durante 2 anos com educação não formal com crianças de 2 a 7 anos (brinquedoteca), há 2 anos atua na Educação Infantil formal com crianças de 3 a 6 anos. Não tem experiência com avaliação ou design de aplicativos.
- Avaliador 5: atua na área de humanas (Pedagogia), tem experiência com crianças de 2 a 7 anos (educação infantil), principalmente com uso de jogos e aplicativos educacionais. Participou de projetos com avaliação de interface de jogos educacionais. Dentre os avaliadores, este foi o único que observou crianças usando os jogos antes de iniciar a sua avaliação.

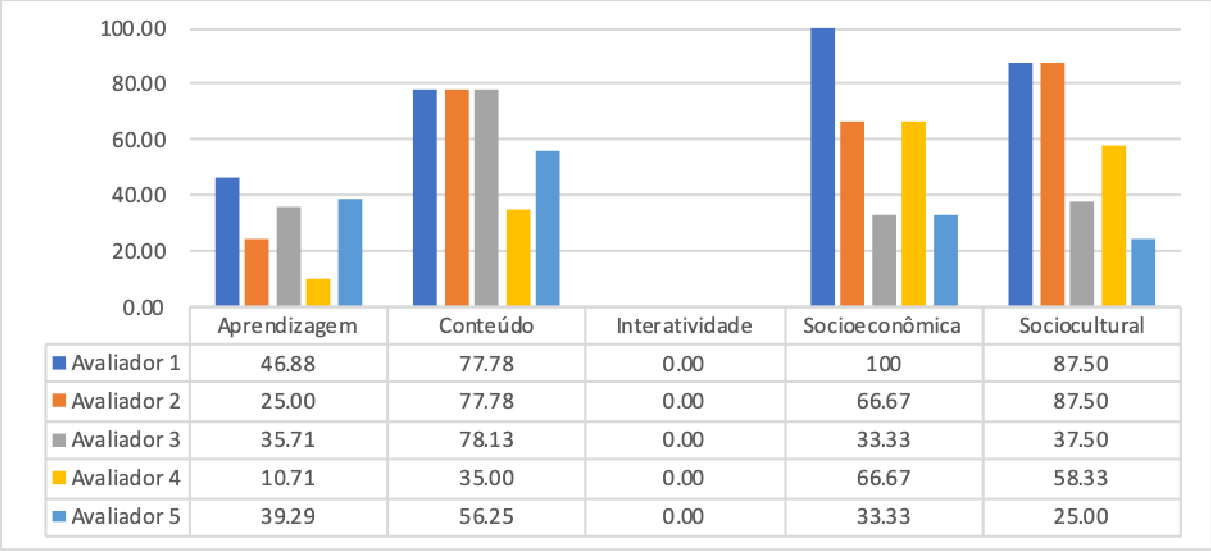
Depois de se familiarizarem com os jogos, os avaliadores utilizaram o questionário final do método MoLEva para caracterizar de forma quantitativa a qualidade dos jogos. Nenhum deles havia realizado avaliações com o método anteriormente. A condução dos testes ocorreu de maneira individual e em diferentes localidades, onde cada avaliador realizou os testes sem nenhuma interferência dos demais.

4. Análise dos Jogos

4.1. Caligrafia

O método MoLEva foi aplicado pelos cinco avaliadores ao jogo Caligrafia. Os Níveis de Qualidade por Critério para cada avaliador são mostrados nas Figura 3 e 4 e os resultados são agrupados na Tabela 3.

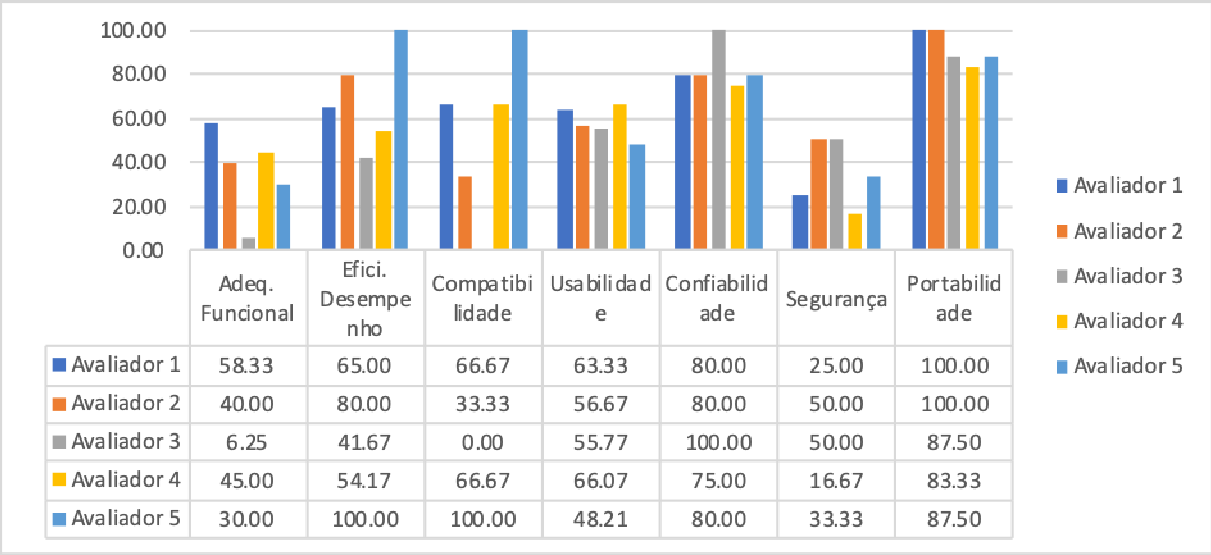
Figura 3. NQRC dos critérios de i a v para o jogo Caligrafia.



	Aprendizagem	Conteúdo	Interatividade	Socioeconômica	Sociocultural
Avaliador 1	46.88	77.78	0.00	100	87.50
Avaliador 2	25.00	77.78	0.00	66.67	87.50
Avaliador 3	35.71	78.13	0.00	33.33	37.50
Avaliador 4	10.71	35.00	0.00	66.67	58.33
Avaliador 5	39.29	56.25	0.00	33.33	25.00

Fonte: Autores

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



	Adeq. Funcional	Efici. Desempenho	Compatibilidade	Usabilidade	Confiabilidade	Segurança	Portabilidade
Avaliador 1	58.33	65.00	66.67	63.33	80.00	25.00	100.00
Avaliador 2	40.00	80.00	33.33	56.67	80.00	50.00	100.00
Avaliador 3	6.25	41.67	0.00	55.77	100.00	50.00	87.50
Avaliador 4	45.00	54.17	66.67	66.07	75.00	16.67	83.33
Avaliador 5	30.00	100.00	100.00	48.21	80.00	33.33	87.50

Fonte: Autores

Tabela 3. Notas dos critérios agrupados em categorias e nota final para o jogo Caligrafia.

CALIGRAFIA					
Critérios/Categorias	**Avaliador 1**	**Avaliador 2**	**Avaliador 3**	**Avaliador 4**	**Avaliador 5**

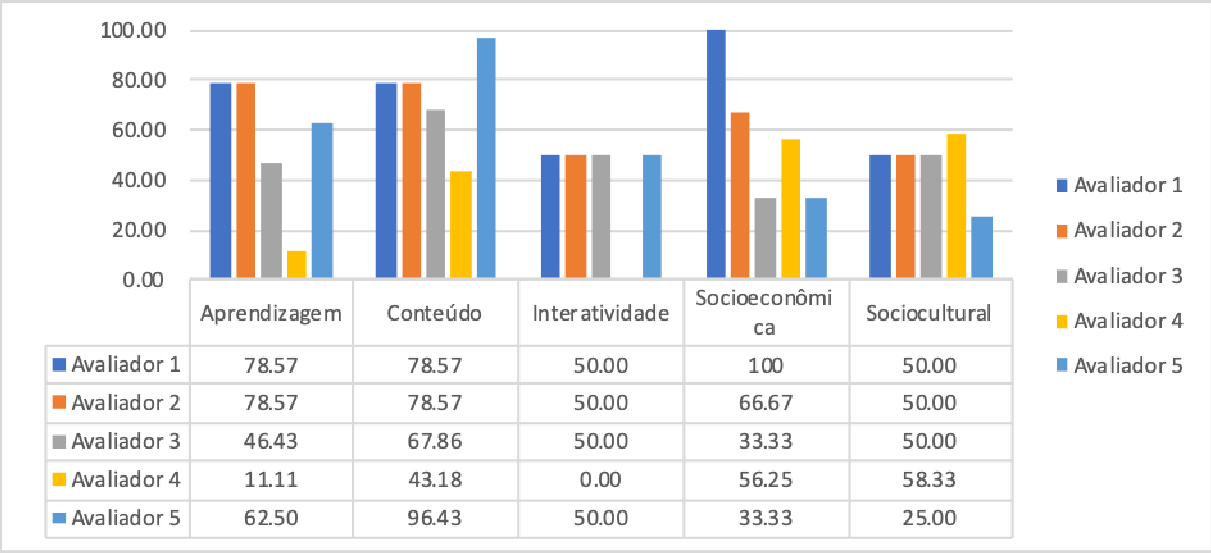
Aprendizagem	46,88	25,00	35,71	10,71	39,29
Conteúdo	77,78	77,78	78,13	35,00	56,25
Interatividade	0,00	0,00	0,00	0,00	0,00
Pedagógica	**41,55**	**34,26**	**37,95**	**15,24**	**31,85**
Socioeconômica	100	66,67	33,33	66,67	33,33
Sociocultural	87,50	87,50	37,50	58,33	25,00
Social	**93,75**	**77,08**	**35,42**	**62,50**	**29,17**
Adeq. Funcional	58,33	40,00	6,25	45,00	30,00
Efici. Desempenho	65,00	80,00	41,67	54,17	100,00
Compatibilidade	66,67	33,33	0,00	66,67	100,00
Usabilidade	63,33	56,67	55,77	66,07	48,21
Confiabilidade	80,00	80,00	100,00	75,00	80,00
Segurança	25,00	50,00	50,00	16,67	33,33
Portabilidade	100,00	100,00	87,50	83,33	87,50
Técnica	**65,48**	**62,86**	**48,74**	**58,13**	**68,44**
NOTA FINAL	**66,93**	**58,07**	**40,70**	**45,29**	**43,15**

Fonte: Autores

4.2. Baú de Brinquedos

De maneira semelhante, o método MoLEva foi aplicado pelos cinco avaliadores ao jogo Baú de Brinquedos. Os Níveis de Qualidade por Critério para cada avaliador são mostrados nas Figura 5 e 6 e os resultados são agrupados na Tabela 4.

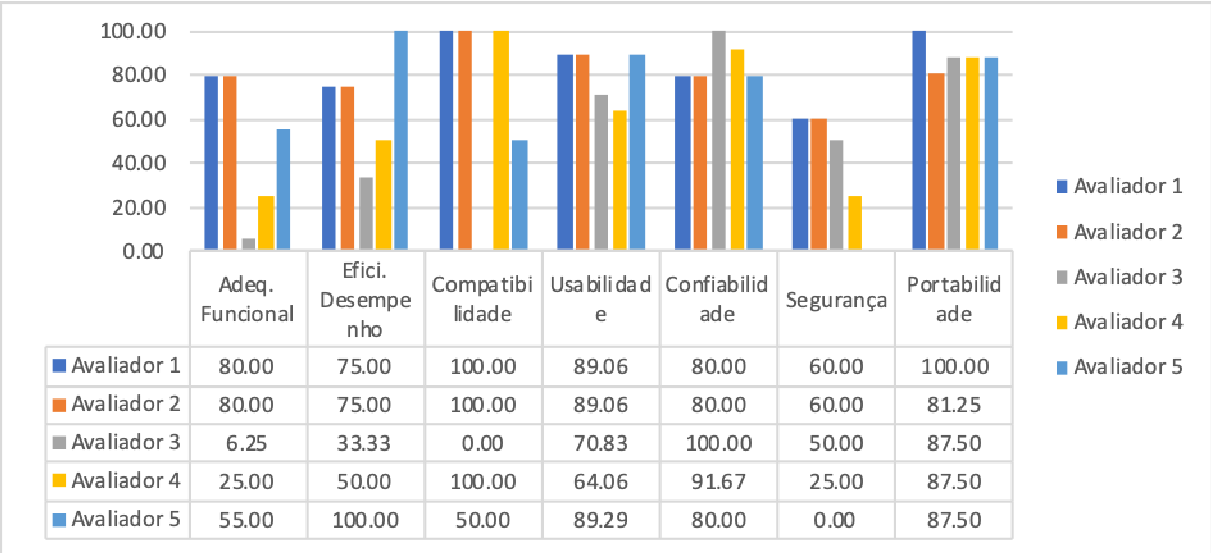
Figura 5. NQRC dos critérios de i a v para o jogo Baú de Brinquedos.



	Aprendizagem	Conteúdo	Interatividade	Socioeconômica	Sociocultural
Avaliador 1	78.57	78.57	50.00	100	50.00
Avaliador 2	78.57	78.57	50.00	66.67	50.00
Avaliador 3	46.43	67.86	50.00	33.33	50.00
Avaliador 4	11.11	43.18	0.00	56.25	58.33
Avaliador 5	62.50	96.43	50.00	33.33	25.00

Fonte: Autores

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



	Adeq. Funcional	Efici. Desempenho	Compatibilidade	Usabilidade	Confiabilidade	Segurança	Portabilidade
Avaliador 1	80.00	75.00	100.00	89.06	80.00	60.00	100.00
Avaliador 2	80.00	75.00	100.00	89.06	80.00	60.00	81.25
Avaliador 3	6.25	33.33	0.00	70.83	100.00	50.00	87.50
Avaliador 4	25.00	50.00	100.00	64.06	91.67	25.00	87.50
Avaliador 5	55.00	100.00	50.00	89.29	80.00	0.00	87.50

Fonte: Autores

Tabela 4. Notas dos critérios agrupados em categorias e nota final para o jogo Baú de Brinquedos.

BAÚ DE BRINQUEDOS

Critérios/Categorias	Avaliador 1	Avaliador 2	Avaliador 3	Avaliador 4	Avaliador 5
Aprendizagem	78,57	78,57	46,43	11,11	62,50
Conteúdo	78,57	78,57	67,86	43,18	96,43
Interatividade	50,00	50,00	50,00	0,00	50,00
Pedagógica	**69,05**	**69,05**	**54,76**	**18,10**	**69,64**
Socioeconômica	100	66,67	33,33	56,25	33,33
Sociocultural	50,00	50,00	50,00	58,33	25,00
Social	**75,00**	**58,33**	**41,67**	**57,29**	**29,17**
Adeq. Funcional	80,00	80,00	6,25	25,00	55,00
Efici. Desempenho	75,00	75,00	33,33	50,00	100,00
Compatibilidade	100,00	100,00	0,00	100,00	50,00
Usabilidade	89,06	89,06	70,83	64,06	89,29
Confiabilidade	80,00	80,00	100,00	91,67	80,00
Segurança	60,00	60,00	50,00	25,00	0,00
Portabilidade	100,00	81,25	87,50	87,50	87,50
Técnica	**83,44**	**80,76**	**49,70**	**63,32**	**65,97**
NOTA FINAL	**75,83**	**69,38**	**48,71**	**46,24**	**54,93**

Fonte: Autores

5. Resultados e Discussão

Nesta seção apresentamos os resultados considerando três frentes: (i) impressões dos Avaliadores 4 e 5 sobre os jogos; (ii) considerações sobre as avaliações considerando o perfil dos avaliadores; (iii) impressões sobre o método MoLEva e implicações para avaliação de jogos educacionais.

5.1 Impressões sobre os jogos analisados

Conforme mencionado na seção 3, o Avaliador 5 observou crianças usando os jogos antes de realizar a sua avaliação. O jogo Caligrafia foi usado por duas crianças, uma de 8 anos e outra de 5 anos. Para ambas, o jogo não se mostrou intuitivo. A criança mais velha, que já sabe escrever, achou que era para passar um risco sobre a linha pontilhada e não fez o desenho da letra, apesar de já saber fazê-lo. A criança mais nova, que ainda não sabe escrever, ficou perguntando o que era para ser feito, selecionou os lápis coloridos e cobriu os pontinhos, sem compreender que era preciso fazer o movimento de desenho da letra. Acabou virando um jogo de ligar pontos.

O Avaliador 4 criticou o Caligrafia, dizendo que realmente não é intuitivo para crianças. De acordo com ele, a criança ao escrever precisa segurar o lápis e fazer um movimento de "pinça". Talvez com o uso de uma caneta digital as crianças tivessem uma compreensão melhor do Caligrafia digital, mas não era o caso em nossos testes, onde a interação foi feita apenas com o dedo.

O jogo Baú de Brinquedos também foi explorado por duas crianças, uma de 3 anos e outra de 5 anos. O Avaliador 5 declarou que não houve muita diferença para as crianças – ambas olharam para o jogo e já compreenderam o que era para ser feito. A única diferença foi que a criança mais nova imediatamente arrastou o objeto até o baú, enquanto a criança mais velha bateu o dedo no objeto, esperando que ele fosse sozinho até o baú, mas logo percebeu que era preciso arrastar com o dedo.

5.2 Considerações sobre as avaliações

As Figuras 3 e 4, bem como a Tabela 3, mostram que para o jogo Caligrafia a avaliação realizada pelos avaliadores 1 e 2 consideram o jogo como nível médio (nota entre 50 e 80). Para os avaliadores 3, 4 e 5 o jogo é de nível inferior (nota menor que 50). O Avaliador 1 não tem nenhuma experiência com crianças, enquanto o Avaliador 2 tem pouca experiência com crianças, quando comparado aos avaliadores 4 e 5. Os únicos com experiência em avaliação/design de interfaces para dispositivos móveis são os avaliadores 3 e 5. Observamos que os avaliadores que têm experiência com crianças e/ou com avaliação de interfaces foram mais rigorosos em suas avaliações.

Vale observar que o avaliador com mais experiencia em avaliação e design de interfaces atribuiu notas inferiores na 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 com relação a esse aspecto.

Ainda no Caligrafia, nota-se que o critério Interatividade foi o único a receber nota zero de todos os avaliadores. Esse critério tem quatro questões que realmente não se aplicam

ao jogo Caligrafia. Os avaliadores selecionaram a opção “Não”, que corresponde a zero pontos e a opção “Não se aplica” que não é computada no cálculo.

No jogo Baú de Brinquedos a avaliação realizada pelos avaliadores 1, 2 e 5 consideram o jogo como nível médio (nota entre 50 e 80). Para os avaliadores 3 e 4 o jogo é de nível inferior (nota menor que 50). Assim como aconteceu no Caligrafia, o Avaliador 3 foi mais rigoroso com relação à categoria Técnica. A maior discrepância é mostrada no critério “Compatibilidade”, onde os avaliadores 1, 2 e 4 atribuíram nota 100 à funcionalidade, o avaliador 3 atribuiu nota 0 e o avaliador 5 nota 50. Isso pode ser explicado uma vez que a maioria das respostas para esse critério tiveram sua avaliação prejudicada, e uma única resposta “Sim” definiu a nota final como 100, pois as demais respostas foram “Não se aplica” ou “Avaliação prejudicada”, que não são computadas na média.

Outro ponto discrepante no jogo Baú de Brinquedos foi em relação à Adequação Funcional. Assim como no caso do Caligrafia, acreditamos que essa discrepância se deve às experiências de cada avaliador. Outra possível explicação é a interpretação dos itens avaliados, uma interpretação incorreta ou divergente pode levar a avaliações diferentes e até mesmo comprometer a avaliação.

É importante observar que, para ambos os jogos, o Avaliador 4 deu notas bem inferiores na categoria “Pedagógica”. 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 ou à prática com crianças. Os avaliadores 4 e 5 têm uma boa experiência com crianças, o que nos leva a questionar por que apenas um deles atribuiu notas tão inferiores à categoria “Pedagógica”. Acreditamos que se trata de uma diferença de perfil. Além de não ter experiência com design e avaliação de aplicativos, o Avaliador 4 não é um adepto às práticas de tecnologias com crianças; tem experiência com crianças e educação infantil, mas tem uma opinião contrária ao uso de tecnologias nessa faixa etária. Por outro lado, o Avaliador 5 tem experiência com avaliação de interface de jogos educacionais e também com crianças no uso de jogos e aplicativos educacionais. Mais uma vez, a nota final de uma categoria foi influenciada pelo perfil dos avaliadores.

5.3 Impressões sobre o método MoLEva

Os avaliadores tiveram dificuldades para responder algumas perguntas por falta de explicação/detalhamento. Por exemplo, na pergunta 43 (Soad, 2017, p. 139) "Como pode ser classificado o consumo de energia do dispositivo?" (com relação ao uso do aplicativo analisado). Para essa pergunta, alguns avaliadores selecionaram a opção "Avaliação prejudicada" e declararam não saber medir esse consumo. O mesmo aconteceu com as perguntas 46, 47 e 84 – "Os requisitos mínimos exigidos pelo aplicativo são respeitados?", "Caso o aplicativo execute funções em segundo plano, essas funções são executadas sem interferirem o desempenho do dispositivo?" e "Como pode ser classificada a facilidade de instalação do aplicativo em todos os dispositivos especificados?". Trata-se de perguntas difíceis de serem respondidas sem uma explicação mais detalhada sobre, por exemplo, quais são os "requisitos mínimos" e quais são os "dispositivos especificados". Diante da dúvida, o avaliador acabava selecionando a opção "Não se aplica" ou "Avaliação prejudicada".

Outra pergunta que gerou dúvida foi a 70, na categoria "Confiabilidade" – "As opções oferecidas pelo aplicativo estão todas funcionando corretamente? Em que nível você as classificaria?" Os avaliadores não sabiam dizer a qual nível a pergunte se refere. Acreditamos que se trata de um erro, pois as perguntas com nível possuem opções entre "Ineficiente" e "Eficiente".

Em alguns casos, os avaliadores compreenderam a pergunta, mas não tinham informações suficientes para respondê-la. Por exemplo, a pergunta 53, da categoria "Usabilidade" - "Os aprendizes com diferentes tipos de experiências/idades têm iguais possibilidades de obter sucesso com as funcionalidades estabelecidas?". Um dos avaliadores, respondeu que "Não" e colocou uma observação dizendo que não tinha informações suficientes para responder. Algo similar aconteceu com a pergunta 66, também da categoria "Usabilidade", sobre "Como pode ser classificada a capacidade do aplicativo em se adaptar a pessoas com necessidades especiais?". Um dos avaliadores assinalou "Avaliação prejudicada" e disse que não testou com pessoas com necessidades especiais, outro avaliador assinalou "Ineficiente" e demostrou dúvida com relação ao tipo de necessidade especial (autista, deficiente físico, visual, auditivo). As perguntas 82 e 83, da categoria "Portabilidade", são sobre a execução do aplicativo em diferentes dispositivos móveis e sistemas operacionais. A avaliação foi prejudicada porque os avaliadores fizeram os testes em apenas um dispositivo.

Com relação ao cálculo da nota final 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 sete critérios enquanto a categoria "Social" possui apenas dois, 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.

Ainda sobre o cálculo da nota, vale observar a diferença entre as respostas "Não", "Não se aplica" e "Avaliação prejudicada". A opção "Não" corresponde a zero pontos, mas esse zero é computado na nota. Já as opções "Avaliação prejudicada" e "Não se aplica" não são computadas na nota, é como se não existissem. Por exemplo, no critério "Compatibilidade" há três perguntas, se o avaliador responder "Não", "Não" e "Sim", a nota final será 33; se responder "Não se aplica", "Não se aplica" e "Sim", a nota final será 100; se responder "Não se aplica", "Não" e "Sim", a nota final será 50. Portanto, as respostas "Não se aplica" e/ou "Avaliação prejudicada" podem levar resultados superficiais ou até mesmo irreais. Isso aconteceu, por exemplo, nas notas do critério "Compatibilidade" no Baú de Brinquedos, conforme citado na seção anterior. A nota geral de três avaliadores foi 100, o que nos leva a pensar que trata-se de um jogo excelente nesse critério; enquanto, na verdade, trata-se de um critério com apenas três perguntas, onde um único "Sim" definiu a nota final como 100, pois as demais respostas foram "Não se aplica" ou "Avaliação prejudicada".

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 das

respostas “Não se aplica” ou “Avaliação prejudicada” pode parecer que resolve o problema em um primeiro momento, mas no final pode gerar uma avaliação superficial, como relatamos no parágrafo anterior.

Por fim, podemos dizer que o método é pouco adequado se o objetivo for avaliar os aplicativos infantis com um olhar mais pedagógico. Uma prova disso é que o jogo Caligrafia foi bastante criticado pelos Avaliadores 4 e 5, teve notas consideradas inferiores na categoria “Pedagógica” por parte de todos os avaliadores. Mas, apesar de tantas críticas aos seus aspectos pedagógicos, o Caligrafia teve uma nota final não muito inferior ao Baú de Brinquedos. Uma sugestão para contornar uma situação como essa, é atribuir um peso maior à categoria “Pedagógica” ou realizar a avaliação considerando apenas essa categoria quando o objetivo da avaliação for analisar esse aspecto.

6. Considerações Finais

Neste artigo apresentamos os resultados da aplicação do método MoLEva para analisar jogos educacionais infantis, considerando diferentes perfis de avaliadores. O resultado ratifica o achado de Batista e colegas (2019) – o perfil do avaliador pode influenciar consideravelmente a avaliação, 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 válida). Observamos que avaliadores que têm experiência com crianças e/ou com avaliação de interfaces foram mais rigorosos em suas avaliações, o que levou a classificações de nível inferior e médio para os mesmos jogos.

Vale observar que, em alguns casos, os avaliadores tiveram dificuldades para responder algumas perguntas por falta de explicação ou por falta de outros dispositivos para realizar os testes, o que pode ter influenciado as notas finais dos aplicativos.

As novas avaliações apresentadas neste artigo, bem como as avaliações em Batista e colegas (2019), nos permitiram tecer impressões mais embasadas a respeito do uso do método para o contexto de jogos educacionais infantis: (i) o uso sucessivo de respostas “Não se aplica” e/ou “Avaliação prejudicada” pode levar a um resultado superficial; (ii) o uso da média aritmética para o cálculo da nota final pode mascarar uma deficiência do aplicativo, caso a nota de uma das categorias tenha sido muito baixa mas as demais tenham sido altas; (iii) o método é pouco adequado se o objetivo for avaliar aplicativos infantis com um olhar

mais pedagógico; (iv) trata-se de um método mais voltado para aspectos técnicos, tanto que essa categoria possui 7 critérios (dentre os 12 totais).

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.

Apesar de se tratar de um método mais voltado para a avaliação de qualidade de aspectos técnicos, o MoLEva pode ser utilizado para avaliação de aspectos pedagógicos com algumas adequações. Nesse caso, sugere-se atribuir um peso maior à categoria "Pedagógica" ou realizar a avaliação considerando apenas essa categoria.

Referências

Amante, L. (2007). "As TIC na Escola e no Jardim de Infância: motivos e factores para a sua integração". Sísifo. Revista de Ciências da Educação, 03, pp. 51‑64. 2007. Disponível em: <http://www.letstry-ict.eu/press/As_TIC_na_Escola_e_no_Jardim_de_Inf__ncia__motivos_e_factores_para_a_sua_integra____o.pdf>. Acesso em: 15 ago. 2019.

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.

Batista, A. C. D.; Melo, L. I. C.; Vascon, L. E. L.; Freire, F. M. P.; Arantes, F. L.; Silva, A. C. 2019. "Avaliação de aplicativos educacionais infantis com o método MoLEva". In: IV Congresso sobre Tecnologias na Educação (Ctrl+E). Recife/PE, Brasil. Agosto de 2019. p. 1-10.

Brasil. Base Nacional Comum Curricular. Brasília: MEC/Secretaria de Educação Básica, 2018. Disponível em: http://basenacionalcomum.mec.gov.br/. Acesso em: 08 out. 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.

Secretaria Municipal de Educação do Estado de São Paulo. 2015. "O uso da tecnologia e da linguagem midiática na Educação Infantil". Diretoria de Orientação Técnica - Educação Infantil. São Paulo, 2015. Disponível em: <portal.sme.prefeitura.sp.gov.br/Portals/1/Files/17138.pdf>. Acesso em: 15 ago. 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. Disponível

em: <https://www.nied.unicamp.br/biblioteca/relatorios-tecnicos/>. Acesso em: 08 out. 2019.

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.

Recebido em Novembro 2019

Aprovado em Novembro 2019