

Orientações para Relato de Pesquisa Quantitativa envolvendo Tecnologias Educacionais

Orientações para Relato de Pesquisa Quantitativa envolvendo Tecnologias Educacionais

1. Introdução

Diversas organizações de apoio à pesquisa científica vêm estabelecendo normas técnicas e recomendações para a publicação em periódicos especializados. A American Psychological Association publicou recentemente dois artigos no periódico *American Psychologist*, descrevendo critérios e normas para apresentação de resultados de pesquisas quantitativas (APPELBAUM et al., 2018) e qualitativas (LEVITT et al., 2018). Da mesma forma, iniciativas semelhantes existem em outras áreas, envolvendo métodos específicos, tais como pesquisas envolvendo neuroimagem (NICHOLS et al., 2017) e ensaios clínicos na medicina (VOHRA et al., 2015). Na grande área de Informática na Educação, temos conhecimento de artigos da comunidade de Inteligência Artificial aplicada à Educação com orientações gerais para condução e relato de avaliações (geralmente experimentais) dos artefatos computacionais resultantes das pesquisas (MARK; GREER, 1993; GREER; MARK, 2016). Essas iniciativas têm dois objetivos principais: (i) garantir padrões mínimos de qualidade e validade das pesquisas realizadas; (ii) padronizar os relatos de pesquisa de modo a facilitar a comparação de resultados e realização de metanálises.

No Brasil, a Comissão Especial de Informática na Educação (CEIE) da Sociedade Brasileira de Computação (SBC) também tem realizado ações semelhantes para a pesquisa em Informática na Educação. Uma dessas iniciativas envolveu o lançamento da série de livros de Metodologia de Pesquisa em Informática na Educação, disponível online e gratuitamente em <https://metodologia.ceie-br.org> (JAQUES et al., 2019). O objetivo da série de livros é atender a pesquisadores e estudantes de pós-graduação em Informática na Educação no Brasil com um material que lhe proporcione um entendimento básico dos fundamentos da metodologia da pesquisa, bem como dos principais métodos e técnicas de coleta e análise de dados. Não menos importante, a série de livros busca fornecer cenários ilustrativos, bem como referencial teórico que permita aos leitores aprofundar os conhecimentos introdutórios apresentados.

Embora a iniciativa proporcione aos pesquisadores brasileiros um entendimento dos fundamentos metodológicos de pesquisa qualitativa e quantitativa, foi identificada a necessidade de documentos de diretrizes que estabelecem padrões para os relatos das pesquisas envolvendo os métodos de pesquisa qualitativa e quantitativa mais frequentes na área de Informática na Educação. O estabelecimento de normas, e o respeito a elas pelos pesquisadores brasileiros, permitirá uma análise e comparação mais rigorosa dos relatos científicos, permitindo verificar e diferenciar achados de artefatos na pesquisa brasileira em Informática na Educação, o que tem sido denominado de pesquisa baseada em evidências (BITTENCOURT; ISOTANI, 2018). Em especial, seguir boas práticas ao descrever os resultados permite o incremento do número de Revisões Sistemáticas com metanálise. Atualmente, essa é uma tendência para classificação das evidências em maior grau (DEMERVAL; COELHO; BITTENCOURT, 2019).

O presente documento sintetiza algumas diretrizes para o relato de publicações científicas e está voltado especificamente às **pesquisas quantitativas**. Na pesquisa quantitativa, os resultados são geralmente obtidos através de dados numéricos e/ou analisados por testes estatísticos. Kumar (2011) apresenta algumas características da

pesquisa quantitativa. Um primeiro aspecto diz respeito ao propósito da investigação, que geralmente se refere à "quantificação da extensão de uma variação em um fenômeno, situação, questão etc" (p.38). Sendo assim, seus tópicos geralmente estão relacionados à "prevalência, incidência, extensão, natureza das questões, opiniões e atitudes", ao invés de, por exemplo, "experiências, significados, percepções e sentimentos" (p. 38). Além disso, com relação à análise de dados, as pesquisas quantitativas tendem a verificar frequência, distribuição, probabilidade e realizar análises estatísticas para testar a magnitude e a força de uma relação entre variáveis. A partir desses resultados, algumas inferências ou conclusões podem ser realizadas.

O presente documento foi desenvolvido por um Grupo de Trabalho formado por pesquisadores brasileiros em Informática na Educação e foi viabilizado por uma ação do Ministério da Educação em parceria com a Universidade Federal de Alagoas e Centro de Inovação para Educação Brasileira (CIEB) no âmbito do projeto Guia de Evidências Educacionais. Os pesquisadores que compõem esse Grupo de Trabalho possuem experiência em pesquisa quantitativa e formação em diversas especialidades (Computação, Neurociência, Psicologia, etc), tendo sido convidados pelo comitê gestor do projeto Guia de Evidências Educacionais.

Este documento foi construído a partir de documentos de entidades internacionais, especialmente o "Artigo de periódico reportando normas para pesquisas quantitativas da American Psychology Association" (APELLBAUM et al., 2018), mas sendo adaptado para o contexto de Informática na Educação e considerando apenas os principais métodos quantitativos empregados na área. Em especial, esse documento apresenta diretrizes para os três métodos principais de pesquisa quantitativa empregados pela comunidade brasileira de pesquisa em Informática na Educação: Estudos experimentais, Estudos correlacionais e Pesquisas de Sujeito Único ou poucos participantes.

Os **estudos experimentais** são empregados por pesquisas que buscam identificar uma relação de causa e efeito (GAZZANIGA; HEATHERTON, 2005), por exemplo, para verificar se o emprego de elementos de gamificação (causa) em tecnologias educacionais leva a mais engajamento (efeito) dos estudantes¹. Um experimento científico consiste na observação de um fenômeno sob condições que o investigador pode controlar. No exemplo dado, o pesquisador vai controlar a presença dos elementos de gamificação, ou seja, um possível experimento seria fazer com que um grupo de participantes usasse, durante um período pré-determinado, um software educacional com elementos de gamificação (grupo experimental) e outro grupo usasse o mesmo software sem gamificação (grupo de controle). Dessa forma, um estudo experimental envolve a seleção aleatória de uma amostra de indivíduos (que serão os participantes do experimento), a distribuição aleatória dos participantes em dois grupos ou mais (controle e experimental) e o emprego de análise estatísticas para inferir conclusões da amostra selecionada para a população. Orientações específicas para relatos de estudos experimentais são apresentadas na Seção 3.1.

Nem sempre, é possível realizar uma observação controlada, por exemplo, para se verificar os possíveis benefícios/mafeícios de políticas públicas educacionais empregadas no passado. Nesse caso, o fenômeno (a política pública em questão) já aconteceu e o que se

¹Os exemplos citados neste documento foram escolhidos para fins didáticos e, por isso, são fáceis de entender e envolvem temas genéricos de conhecimento da comunidade científica. Assim, embora eles sigam as recomendações técnicas, eles não serviriam como temas atuais de pesquisa porque não possuem originalidade científica. Embora não abordado neste capítulo, essa é uma característica central de uma pesquisa e deve ser levada em conta pelo pesquisador na definição do seu trabalho de pesquisa.

pode fazer nesse sentido é verificar se existe uma associação entre os fenômenos e consequências esperadas (maior retenção escolar, menos desistência, etc). Nessa situação, temos um **estudo correlacional**, um tipo de pesquisa descritiva, em que se busca verificar se existe uma correlação entre um fenômeno e outras variáveis. Em um estudo correlacional não se pode afirmar que existe uma relação de causa e efeito (por exemplo, pode haver uma terceira variável que impacta nas duas estudadas), mas apenas que existe uma associação entre duas variáveis. Na Seção 3.2, são introduzidas algumas orientações específicas para estudos correlacionais.

Uma outra variação para os estudos experimentais são os trabalhos denominados como **“sujeito único”** ou **“poucos participantes”** [small-N]. Nesse tipo de trabalho, o conjunto de dados não é suficiente para a utilização de métodos estatísticos tradicionais, entretanto, o delineamento do experimento é planejado de modo a garantir controle experimental rigoroso. A lógica básica é a comparação dos resultados de cada participante com seus próprios dados, acompanhando como o tratamento afetou seu desempenho. Para aumentar a probabilidade de que as alterações são geradas pelo tratamento, são utilizadas estratégias como reversão, linha de base múltipla, sondas alternadas etc. Essas estratégias são brevemente discutidas na Seção 3.3.

Este documento está organizado da seguinte forma. Na Seção 2, são apresentadas diretrizes gerais para artigos quantitativos, tais como sugestões de como definir títulos, escrever resumos e introduções e orientações gerais para as seções de métodos, análise e discussão dos resultados. Nas próximas seções, descrevemos diretrizes específicas para guiar os três métodos de pesquisa quantitativos mais empregados nos trabalhos brasileiros de pesquisa em informática na educação: estudos experimentais (Seção 3.1), estudos correlacionais (Seção 3.2) e pesquisas de sujeito único ou poucos participantes (Seção 3.3).

Estudo de Caso não é um método de pesquisa quantitativo

Existem algumas definições sobre o que vem a ser um "estudo de caso". Neste documento, consideramos "estudo de caso" a investigação minuciosa de unidades individuais (como um aluno, uma escola ou um grupo particular de indivíduos) a partir de um prisma predominantemente da **pesquisa qualitativa**. Sendo assim, este quadro tem por objetivo esclarecer que **estudos quantitativos associados a amostras pequenas** ($n < 50$), quando há uma abordagem quantitativa em seu escopo e resultados, não serão considerados como estudo de caso.

Em publicações nacionais que fazem uso de tecnologia na educação, é bem comum realizar um experimento dentro de uma determinada escola ou em um contexto de participantes bem limitado, e no título do trabalho é observado que os autores classificaram a pesquisa como sendo estudo de caso. Se os dados da pesquisa tiverem sido coletados a partir de observações e entrevistas, estruturadas ou semiestruturadas, as quais são abordagens bem comuns em pesquisa qualitativa, então o enquadramento está correto. No entanto, se no estudo há, por exemplo, a comparação de métricas como desempenho, engajamento, além da distinção clássica entre grupo de controle e experimental, e outras abordagens oriundas da pesquisa quantitativa, então este não será considerado um estudo de caso, mas sim como pesquisas quantitativas de sujeito único ou poucos participantes. Para uma melhor compreensão sobre os estudos experimentais envolvendo poucos participantes, sugere-se a leitura da Seção 3.3.

2. Orientações Gerais

Para iniciar, nessa seção vamos apresentar algumas diretrizes para guiar a escrita de artigos que seguem algum método de pesquisa quantitativo. Essa seção está dividida em título, resumo, introdução, método, resultados e discussão. Esta é a organização geral recomendada pela APA para escrita de artigos quantitativos. Assim como nas outras seções deste artigo, nos baseamos fortemente nas orientações do artigo de Appelbaum et al. (2018) para as sugestões descritas nessa seção, uma vez que a comunidade científica americana em Psicologia tem um histórico reconhecido internacionalmente na definição dessas diretrizes, principalmente para estudos quantitativos.

No **Título**, sugere-se a inclusão das principais variáveis de trabalho (independente, dependentes e outras, se houver), da relação entre elas, assim como das principais questões teóricas sob investigação. É também recomendado incluir a população estudada.

O **Resumo** deve descrever os objetivos do trabalho, participantes, método quantitativo empregado, assim como resultados e suas implicações. O *resumo não deve apresentar referências*, exceto se estas foram fundamentais para a compreensão do trabalho, como no caso de uma replicação ou uma crítica específica a um artigo. Karumbaiah et al. (2018) é um exemplo dessa segunda situação.

A **Introdução** deve apresentar o problema, os objetivos e a justificativa para o estudo, incluindo o contexto para entendê-lo, bem como sua relevância. Uma revisão da literatura, sucinta, mas que permita identificar o referencial do trabalho proposto e a lacuna no conhecimento a ser explorada também devem ser apresentada. As pesquisas quantitativas geralmente são explicativas e, por isso, envolvem hipóteses a serem testadas. Essas hipóteses devem ser explicadas sobre a luz de uma teoria. Por exemplo, se a hipótese de uma pesquisa é que o *ranking* (elemento de gamificação) deve engajar mais os alunos com traço de personalidade extroversão, são necessárias explicações plausíveis da literatura científica baseada em evidências que justifiquem essa hipótese. Assim, ao fim da introdução, sugere-se apresentar as hipóteses da pesquisa e uma breve justificativa.

A seção do Método (**ou Metodologia, ou Materiais e Métodos**) é um elemento crucial de um artigo, na qual o pesquisador deve fornecer informações sobre como a mesma foi conduzida, seus delineamentos e desdobramentos até chegar aos resultados. Esta seção é importante para garantir a reprodutibilidade da pesquisa e também permitir a avaliação da qualidade dos protocolos e técnicas empregados. Nesta seção, recomenda-se inicialmente descrever as características dos participantes do estudo (por ex., gênero, idade, etnicidade, escolaridade, status socioeconômico, etc). A forma como a amostra foi coletada também precisa estar descrita, de modo a fornecer ao leitor como ocorreu o processo de amostragem e recrutamento. O artigo deve informar o tamanho de amostra pretendido a priori, o que foi efetivamente obtido e as causas da exclusão de indivíduos. Além do tamanho, recomenda-se algumas medidas para verificar se o tamanho da amostra é suficiente - por exemplo, poder da estatística. A aprovação do Comitê de Ética local para condução do estudo, bem como o consentimento livre e esclarecido dos participantes também devem ser descritos nesta seção.

Em seguida, o pesquisador deve descrever o método empregado (ex: protocolos, materiais, sistemas, etc), delineamento (ex: experimental, correlacional etc), procedimentos e métodos de análise estatística (incluindo o nível de significância a ser considerado e os métodos para correção por múltiplas comparações). Reforça-se que o nível de detalhamento destas descrições deve ser o suficiente para assegurar a reprodutibilidade do estudo por qualquer pesquisador independente.

Sugere-se seções separadas para Resultados e Discussão. Na seção de **Resultados**, os achados das análises estatísticas realizadas e os resultados são apenas descritos (e os principais enfatizados), mas evitando qualquer tipo de argumentação. Já na seção de **Discussão**, esses resultados devem ser discutidos sob a luz das teorias na qual o trabalho se fundamenta. Os autores devem buscar explicar se as hipóteses são confirmadas ou não pelos resultados, e em caso negativo, buscar possíveis explicações. Na Discussão, os autores também devem procurar situar os seus achados em relação ao estado da arte: o que foi encontrado que é suportado por outras pesquisas ou contradiz outros resultados científicos. Finalmente, deve ser apresentada uma discussão das implicações dos resultados para pesquisas futuras.

O Quadro 1 sumariza as principais orientações para os estudos científicos quantitativos em geral, independente do método empregado por ele.

Quadro 1 - Orientações Gerais para Relatos de Estudos Científicos Quantitativos

Título	Identifique as principais variáveis do trabalho (independente, dependente, etc), bem como a população investigada. Indique os conceitos teóricos mais relevantes para o trabalho.
Resumo	<p>Objetivos</p> <ul style="list-style-type: none"> - Declare o problema a ser investigado; - Declare suas hipóteses; <p>Participantes</p> <ul style="list-style-type: none"> - Descreva os participantes, especificando suas características pertinentes ao estudo. <p>Método</p> <ul style="list-style-type: none"> - Descreva o método do estudo, incluindo: - Delineamento da pesquisa (por exemplo, experimental, correlacional, etc); - Tamanho da amostra; - Materiais utilizados (incluindo instrumentos e aparatos); - Medidas empregadas para apresentar os resultados do trabalho; - Procedimentos de coleta de dados; <p>Resultados</p> <ul style="list-style-type: none"> - Descreva os resultados, incluindo tamanho de efeito (por exemplo, d de Cohen), intervalo de confiança e p-valor. <p>Conclusões</p> <ul style="list-style-type: none"> - Apresente conclusões, indo além dos resultados, de modo a relatar suas implicações ou aplicações.
Introdução	Problema

	<ul style="list-style-type: none"> - Apresente a importância do problema, incluindo implicações práticas e teóricas que justifiquem o estudo; <p>Revisão de literatura</p> <ul style="list-style-type: none"> - Ofereça uma revisão sucinta da literatura, incluindo: - Relações do trabalho atual com estudos anteriores; - Diferenças entre o trabalho atual e publicações anteriores, caso algum aspecto do estudo tenha sido reportado anteriormente; <p>Hipóteses, propostas e objetivos</p> <ul style="list-style-type: none"> - Apresente qualquer hipótese específica, propostas e objetivos, incluindo: - Teorias ou outros meios que foram utilizados para estabelecer a hipótese; - Hipóteses primárias e secundárias; - Declare como as hipóteses e o delineamento da pesquisa estão relacionados.
Método	<p>Características dos participantes</p> <ul style="list-style-type: none"> - Descreva características demográficas gerais (por exemplo, idade, sexo, etnia, status socioeconômico); <p>Inclusão e exclusão</p> <ul style="list-style-type: none"> - Descreva os critérios de inclusão e exclusão, incluindo quaisquer restrições baseadas em características demográficas; <p>Procedimentos de amostragem</p> <ul style="list-style-type: none"> - Descreva os procedimentos para seleção de participantes, incluindo: - Método de recrutamento; - Método de amostragem, se um método sistemático foi implementado. - Porcentagem da amostra de participantes que realmente participou; - Se ocorreu auto-seleção (seja por indivíduos, seja por unidades, como escolas ou clínicas); - Materiais e locais onde os dados foram coletados, assim como datas da coleta de dados; - Procedimentos para consentimento livre e esclarecido por parte dos participantes; - Número do protocolo do comitê de ética e descrição dos cuidados éticos realizados. <p>Tamanho e poder [power] da amostra</p> <ul style="list-style-type: none"> - Descreva o tamanho da amostra, poder e precisão, incluindo: - Tamanho de amostra pretendido; - Tamanho de amostra atingido, se diferente do pretendido; - Determinação do tamanho da amostra, incluindo: - Poder de análise, ou métodos utilizados para determinar precisão dos parâmetros estimados; <p>Tipos de Medidas e Variáveis</p> <ul style="list-style-type: none"> - Descreva os tipos de variáveis que serão observadas, a sua natureza e seu nível de medida;

- Defina todas as medidas primárias e secundárias, incluindo medidas coletadas, mas não incluídas no trabalho.

Coleta de dados

- Descreva métodos usados para coletar dados.

Qualidade das medidas

Tão importante quanto a coleta de dados é a fidedignidade da informação associada aos dados coletados. Por isso:

- Descreva métodos usados para melhorar a qualidade das medidas, incluindo:

- Treinamento e confiabilidade dos coletores de dados;
- Uso de múltiplos observadores.

Instrumentos

- Ofereça informação sobre instrumentos validados ou *ad hoc* criados para estudos individuais, por exemplo, propriedades psicométricas e biométricas

Estudo cego ou duplo cego

- Relate se os experimentadores, sejam aqueles administrando as manipulações experimentais ou aqueles avaliando os resultados da pesquisa, conhecem as condições avaliadas.
- Se o procedimento cego for utilizado, explique como ele foi realizado e que medidas de avaliação foram tomadas para avaliar se este foi bem-sucedido.

Condições e delineamento

Os tipos de pesquisas descritas neste protocolo deverão ser levados em consideração para o preenchimento desta seção, então:

- Declare se as condições foram manipuladas ou naturalmente observadas. Relate o tipo de delineamento da pesquisa: Manipulação experimental com participantes randomizados, Manipulação experimental sem randomização, Delineamento não-experimental (sem manipulação experimental), Delineamento longitudinal, Estudos de sujeito único, Replicações.

Diagnóstico dos dados

- Descreva o diagnóstico de dados, incluindo:
 - Critério para exclusão de participantes após a coleta de dados, se houver;
 - Critério para decidir se haverá inferência de dados perdidos [missing data] e métodos usados para imputação de dados perdidos;
 - Definição e processamento estatístico de *outliers*;
 - Análises da distribuição de dados;
 - Transformações nos dados que serão usadas, se houver;

Estratégia de Análise Estatística

- Descreva qual método estatístico será utilizado para a análise de dados. Este é fundamental para garantir a validade do que será apresentado nos resultados. Tais modelos precisam ir ao encontro das hipóteses geradas, tendo que ser capazes de confirmá-las, ou não.

	<ul style="list-style-type: none"> - Descreva o nível de significância (Erro Tipo I) a ser considerado (ex: 5%); - Descreva o método de correção para comparações múltiplas aplicado (ex: método de Bonferroni). A aplicação dessas correções é fundamental para evitar a inflação do Erro do Tipo I e evitar falsos positivos.
Resultados	<p>Fluxo de participantes</p> <ul style="list-style-type: none"> - Relate o fluxo de participantes, incluindo número total de participantes em cada grupo para cada estágio do estudo; <p>Estatísticas e análises de dados</p> <p>Ofereça informações detalhando os métodos estatísticos e de análise de dados empregados, incluindo:</p> <ul style="list-style-type: none"> - Dados perdidos, incluindo: Frequência ou porcentagem de dados perdidos, acompanhados de evidência empírica e/ou argumentos teóricos que avaliem as razões da perda; e Métodos empregados para lidar com dados perdidos, se houver; - Descrição de cada resultado em termos de número de casos, média, desvio padrão e outras medidas que caracterizam os dados empregados; - Estatística inferencial, incluindo: <ul style="list-style-type: none"> - Resultados de todos os testes inferenciais conduzidos e os métodos que foram empregados para avaliar valores de p exatos na testagem estatística de hipótese nula; - Estimativas de tamanho do efeito e intervalos de confiança referentes a cada teste inferencial conduzido, quando possível; - Identificação do software estatístico usado para realizar a análise (por exemplo, SAS PROC GLM, ou uma biblioteca R particular), incluindo sua versão; - Descreva quaisquer problemas com suposições de distribuição de dados que possa afetar a validade dos achados;
Discussão	<p>Suporte da hipótese original</p> <ul style="list-style-type: none"> - Declare se houve suporte para as hipóteses; <p>Similaridade dos resultados</p> <ul style="list-style-type: none"> - Discuta similaridades e diferenças entre os resultados relatados e os trabalhos relacionados; <p>Interpretação</p> <ul style="list-style-type: none"> - Ofereça uma interpretação dos resultados, levando em consideração: <ul style="list-style-type: none"> - Fontes de vieses potenciais e ameaças à validade interna e estatística; - Imprecisão nos protocolos de medidas; - Adequação do tamanho da amostra e validade da amostra; <p>Generabilidade</p> <ul style="list-style-type: none"> - Discussão da generabilidade (validade externa) dos achados, levando em consideração: <ul style="list-style-type: none"> - População alvo (validade de amostra);

	<ul style="list-style-type: none"> - Outras questões contextuais (ambiente, medidas, tempo, validade ecológica); <p>Implicações</p> <ul style="list-style-type: none"> - Discuta implicações para pesquisas futuras, programas ou políticas. <p>Limitações</p> <ul style="list-style-type: none"> - Discussão das possíveis limitações e fontes de erro que não puderam ser controladas;
--	---

Fonte: Adaptado de Appelbaum et al. (2018).

3. Tipos de Pesquisa

Nessa seção são descritas as diretrizes específicas para os três tipos principais de métodos de pesquisa quantitativos empregados pela comunidade brasileira de pesquisa em Informática na Educação: estudos experimentais, estudos correlacionais, e estudos quantitativos envolvendo grupos pequenos.

Embora haja outros métodos quantitativos que são descritos em outros manuais e documentos de diretrizes, por exemplo, pesquisas baseadas em redes bayesianas, incluída no relatório da APA (APPELBAUM et al., 2018), escolhemos por incluir neste documento apenas os métodos quantitativos mais usuais nas pesquisas de Informática na Educação publicadas nos veículos da CEIE (RBIE, SBIE, WIE, etc), para que esse documento ficasse conciso. Dessa forma, recomendamos o documento da APA para outros métodos quantitativos mais específicos não incluídos neste capítulo.

3.1 Estudos Experimentais

Um método de pesquisa comumente empregado em informática na educação são os estudos experimentais ou delineamentos experimentais, que são utilizados quando desejamos avaliar o efeito de um software educacional ou estratégia pedagógica nos estudantes. No experimento, é importante definir claramente a(s) variável(is) independente(s) (o que se deseja manipular, verificar o efeito), a(s) variável(is) dependente(s) (que depende da variável independente) e o efeito em si (melhora, piora, etc). Os experimentos envolvem a manipulação da variável independente de forma a se verificar um efeito na variável dependente. Portanto, há **manipulação da variável independente**.

Quando se realiza um experimento, tem-se geralmente uma, ou mais, hipótese(s) de qual vai ser o efeito da(s) variável(is) independente(s) na(s) dependente(s). Essa hipótese usualmente é definida a partir de estudos anteriores e teorias, e deseja-se verificar se essa hipótese se sustenta, por meio de um experimento.

Como exemplo, suponha que um pesquisador acredita que o emprego de gestos por agentes pedagógicos animados em sistemas tutores inteligentes (para salientar partes importantes de

uma equação) melhora a compreensão dos estudantes sobre resolução de equações algébricas. O pesquisador então irá desenhar um experimento em que uma amostra de estudantes será dividida em dois grupos: um grupo aprenderá resolução de equações no sistema tutor com um agente que não emprega gestos para ensinar (geralmente, chamado de **grupo de controle**) e o outro grupo usará uma outra versão do tutor em que os agentes animados usam gestos para salientar partes da equação durante a sua explicação (**grupo experimental**). A presença de um grupo de controle é importante para podermos verificar se a aprendizagem melhora pela presença de gestos do agente animado ou por algum outro fator (por exemplo, a simples presença do agente). Por isso, o grupo de controle deve ser o mais semelhante possível ao grupo experimental, tendo como principal (desejavelmente única) diferença a variável independente. Desse modo, se for encontrado um efeito, a causa mais provável é a variável independente, já que as condições em ambos os grupos foram "controladas" para serem as mais similares possíveis.

Uma questão importante é como esses estudantes serão distribuídos para compor os dois grupos. Idealmente, esses estudantes devem ser distribuídos aleatoriamente, **randomização**, para que os participantes tenham chances iguais de serem designados para qualquer um dos grupos, aumentando as chances dos grupos serem equivalentes em termos de características dos participantes do experimento. Alternativamente, é possível avaliar o desempenho inicial dos participantes em relação ao maior número possível de variáveis relevantes e, então, fazer um balanceamento para homogeneizar os grupos. Em alguns casos, os dois métodos são combinados, com distribuição aleatória e posterior balanceamento a partir de avaliações iniciais. Dessa forma, se um resultado for encontrado (se ocorreu o efeito) e se a hipótese nula (não há diferença entre os grupos) é falsa; a possível causa é a manipulação experimental e não as características dos participantes. Por outro lado, imagine a situação em que um experimentador recrutou estudantes na rede social da sua universidade e designou os primeiros 50 participantes ao grupo experimental e os últimos 50 ao grupo de controle. Uma outra possível explicação para o efeito encontrado na variável dependente é que possivelmente os estudantes do grupo experimental já sejam mais naturalmente engajados para participar do experimento e, por isso, chegaram primeiramente.

Nesta seção, serão descritas algumas orientações, no Quadro 2, para auxiliar pesquisadores a escreverem as seções de Método, Resultados e Discussão de um artigo, relatório ou monografia envolvendo estudos experimentais, sejam com distribuição aleatória ou não dos participantes. Além das informações solicitadas na Seção 2, você deve também dar detalhes sobre as manipulações experimentais contidas nas orientações de Método do Quadro 2. Na seção Resultados e Discussão, além de descrever os aspectos indicados na Seção 2 (Quadro 1), sugere-se também seguir as orientações do Quadro 2, específicas para estudos experimentais.

Quadro 2 - Orientações Específicas para Relatos de Estudos Experimentais

Método	- Apresente detalhes das manipulações experimentais pretendidas para cada condição de estudo, incluindo condições de comparação, e como e
---------------	---

	<p>quando manipulações experimentais foram atualmente administradas, incluindo:</p> <ul style="list-style-type: none"> - Conteúdos das manipulações experimentais específicas; - Sumário ou paráfrase de instruções utilizadas. Caso estas sejam incomuns ou façam parte das manipulações experimentais, elas devem ser apresentadas de modo literal e completo. - Método de realização da manipulação experimental; <ul style="list-style-type: none"> - Descrição de aparelhos e materiais utilizados e sua função no experimento; - Experimentador: quem introduziu as manipulações experimentais, especificando sua formação profissional e treinamento para realização dos procedimentos específicos; - Número de experimentadores, - Relate se o experimento é cego, duplo cego ou outro: os participantes e experimentadores que administraram as manipulações experimentais, e avaliaram os resultados, estavam cientes das condições da avaliação; - Ambiente: onde as intervenções ou manipulações experimentais ocorreram; - Duração e quantidade da exposição: quantas sessões, episódios ou eventos foram planejados e realizados e duração; - Intervalo de tempo: quanto tempo levou para realizar a manipulação experimental para cada unidade; - Atividades para aumentar aderência ou conformidade (por exemplo, incentivos);² <p>É importante ter em mente que deve ser dado detalhamento suficiente para permitir a replicação do experimento. Se necessário, pode ser incluído uma cópia ou link de um manual de procedimento.</p> <p>Participantes</p> <ul style="list-style-type: none"> - Informe o número inicial de participantes em cada grupo e como eles foram designados para seus respectivos grupos (distribuição aleatória ou não e como ela foi realizada); - Observação: na maioria dos experimentos, a unidade de análise são os indivíduos (participantes), por isso, nesse documento, para fins didáticos, estamos considerando que essa é a unidade de análise. No entanto, pode haver experimentos em que se deseja comparar escolas ou grupos de trabalho e isso deve ser claramente documentado na seção de Métodos.
Resultados	Fluxo de participantes

² No Brasil, ao contrário de outros países, não é possível realizar pagamentos diretos a participantes de pesquisa, entretanto, é possível ressarcir-los de custos eventuais com transporte e/ou alimentação. Além disso, alguns pesquisadores arranjam situações que preveem algum tipo de benefício relacionado à participação na pesquisa, como acesso a devolutivas de resultados, intervenções relevantes para os participantes etc. Caso essas estratégias tenham sido utilizadas, é importante relatá-las.

	<ul style="list-style-type: none"> - Relate o número total de grupos (se manipulações experimentais que foram administrados ao nível do grupo) e o número de participantes assinalados para cada grupo, incluindo: <ul style="list-style-type: none"> - Número de participantes abordados para inclusão; - Número de participantes que iniciaram o experimento; - Número de participantes que não completaram o experimento ou que transitaram entre condições, acompanhado das razões para isso; - Número de participantes incluídos; <p>Fidedignidade do tratamento</p> <ul style="list-style-type: none"> - Ofereça evidência se as manipulações experimentais foram implementadas conforme planejado; se houver algum tipo de procedimento para verificar a integridade do procedimento, relatar. <p>Eventos adversos e efeitos colaterais</p> <ul style="list-style-type: none"> - Relate todos os eventos adversos ou eventos colaterais de cada condição experimental. Se não houver, declare isso.
Discussão	<ul style="list-style-type: none"> - Discuta os resultados, levando em consideração o mecanismo pelo qual a manipulação experimental teria funcionado (relações causais) ou mecanismos alternativos; - Discuta o sucesso de, ou obstáculos à, implementação das manipulações experimentais; - Discuta generalidade (validade externa e validade de constructo) dos achados, levando em consideração: <ul style="list-style-type: none"> - Características da manipulação experimental; - Como, quais resultados foram medidos; - Incentivos; - Descreva a significância teórica ou prática dos resultados e a base para essas interpretações.

Fonte: Adaptado de Appelbaum et al. (2018).

3.2 Estudo Correlacional

Além dos estudos experimentais, há outros métodos para investigar e avaliar a eficácia de tecnologias educacionais. Entre estes, os estudos correlacionais vêm se demonstrando especialmente relevantes com o desenvolvimento da tecnologia, uma vez que uma grande quantidade de dados passou a ser registrada em situações cotidianas. A principal diferença entre um estudo correlacional para um experimental é que no primeiro não há manipulação das condições, sendo predominantemente observacional (isto é, sem a aplicação de uma intervenção). Sendo assim, atualmente, é possível acessar dados sobre as mais diferentes temáticas, incluindo registros escolares, prontuários médicos, transações comerciais e até mesmo interações em redes sociais. Adicionalmente, potentes ferramentas de mineração e análise de dados vêm sendo desenvolvidas para identificar e investigar relações entre variáveis e construir modelos explicativos para os mais diferentes fenômenos.

Cabe registrar que a pesquisa correlacional tem origem anterior a dos desenvolvimentos tecnológicos que geraram o processo usualmente denominado como *big data*, mas sua relevância aumentou consideravelmente a partir deste. Pesquisas correlacionais são estudos que aplicam métodos estatísticos para análise de associações entre variáveis ou avaliação de poder preditivo de variáveis em relação a outras. Ela inclui, portanto, os *surveys* (pesquisas de opinião), os estudos de dados referentes à políticas públicas, como, por exemplo, avaliações das escolas públicas, prontuários do SUS etc, e qualquer outro trabalho que obtenha dados fora de situações experimentalmente controladas.

Sua primeira relevância é exploratória e descritiva, uma vez que estudos correlacionais permitem conhecer características de grandes populações ou compreender globalmente fenômenos complexos. Além disso, os estudos correlacionais utilizam metodologias para avaliar a relações entre variáveis, bem como para constituir modelos explicativos a partir de um conjunto de dados a respeito de um fenômeno.

Nessa seção, serão abordados os principais aspectos a serem considerados em um estudo correlacional em adição àqueles discutidos nas diretrizes gerais. Essas orientações específicas são apresentadas no Quadro 3.

Quadro 3 - Orientações Específicas para Relatos de Estudos Correlacionais

<p>Método</p>	<p>Seleção de participantes</p> <ul style="list-style-type: none"> - Descreva o(s) método(s) de seleção de participantes (por exemplo, as unidades a serem observadas, classificadas etc), incluindo: <ul style="list-style-type: none"> - Método(s) de seleção dos participantes de diferentes grupos (por exemplo, métodos de amostra, local do recrutamento etc) e o número de casos em cada grupo; - Critério de pareamento entre grupos, caso algum tipo de pareamento tenha sido utilizado; - Identifique fontes de dados usadas (por exemplo, fontes de observação, registros de arquivos) e, se relevante, inclua códigos ou algoritmos para selecionar os registros relacionados a cada participante; <p>Variáveis</p> <ul style="list-style-type: none"> - Defina todas as variáveis claramente, incluindo: <ul style="list-style-type: none"> - Exposição; - Preditores potenciais, confundidores e modificadores de efeito; - Explique como cada variável foi medida (e o tipo de cada variável, ex: categórica nominal, ordinal, quantitativa discreta ou contínua); <p>Possibilidade de comparar as avaliações</p> <ul style="list-style-type: none"> - Descreva a comparabilidade da avaliação entre grupos (por exemplo, a probabilidade de observação ou registro de um resultado em um grupo por razões não relacionadas ao efeito da intervenção); <p>Análise</p> <ul style="list-style-type: none"> - Descreva como preditores, fatores que geram confusão e modificadores de efeito foram incluídos na análise;
<p>Discussão</p>	<p>Limitações na discussão</p> <ul style="list-style-type: none"> - Descreva limitações potenciais do estudo.

Fonte: Adaptado de Appelbaum et al. (2018).

3.3 Pesquisas de Sujeito Único ou poucos participantes [Small n]

As pesquisas de sujeito único ou poucos participantes são um tipo específico de pesquisa experimental que não utiliza o delineamento de grupos controle e experimental³. Esse modelo de estudo surgiu no contexto da educação inclusiva, para a identificação de tratamentos eficazes para pessoas que apresentavam transtornos específicos e, frequentemente, raros. Nesses casos, a constituição de grupos homogêneos com grande quantidade de participantes era muito difícil, de modo que se tornou necessário o desenvolvimento de métodos de controle experimental que pudessem ser aplicados para um único caso ou para poucos casos. Como nos estudos experimentais clássicos, as pesquisas de sujeito único ou pequenos grupos são planejadas de modo a garantir validade interna e externa. A validade interna avalia se os resultados verificados em um estudo específico têm generalidade para outros participantes, ambientes ou comportamentos. A validade externa verifica se as mudanças observadas são uma função da variável independente e não de variáveis intervenientes. Ambas são obtidas por meio da comparação dos resultados em diferentes condições, ainda que com o mesmo participante ou com poucos participantes (COOPER; HERON; HEWARD, 1987). A seguir, são descritos os principais delineamentos utilizados nesse tipo de pesquisa.

Delineamento A-B

Usado para comparação das condições antes e depois do tratamento. Para isso, é necessário estabelecer uma situação controle, ou linha de base inicial (A), seguida pela condição experimental (B). O efeito é demonstrado por meio da mudança nos resultados de uma fase para a outra. Então, por exemplo, na fase de linha de base (medidas iniciais), pode-se verificar a quantidade de tempo que um estudante passa em determinado aplicativo educacional e, posteriormente, a mesma medida é realizada após a realização de mudanças no aplicativo. Importante observar que a efetividade da comparação vai depender da confiabilidade da medida inicial. Portanto, é fundamental que as medidas sejam tomadas até que se estabeleça uma linha de base estável, ou seja, que circunscreva as possíveis variações observadas. A quantidade de tempo necessária para essa estabilidade vai variar amplamente dependendo do que está sendo medido. No exemplo citado, usar apenas alguns dias pode não oferecer uma medida segura porque o estudante pode estar passando por um período de provas, ou estar doente etc. Também pode haver uma alteração cíclica que precisa ser adequadamente registrada, por exemplo, o aluno passa bastante tempo no aplicativo na segunda e na terça-feiras e pouco tempo no restante dos dias da semana.

Após a implementação das mudanças que são o alvo do trabalho, a mesma medida deve ser realizada até que se verifique estabilidade. Os resultados são, então, comparados. Caso as alterações no aplicativo estejam relacionadas com o tempo de uso do mesmo, espera-se observar um aumento no tempo total ou na frequência de utilização da ferramenta.

O delineamento A-B é fácil de realizar e apresenta-se como uma primeira forma de avaliar resultados, entretanto, ele tem limitações importantes. Em primeiro lugar, não há avaliação da replicabilidade dos resultados, de modo que não é possível afirmar que a mesma mudança

³ Apesar de esse tipo de investigação ser usualmente denominada como “sujeito único” [single subject] e grupos pequenos [small n], não é necessário que o trabalho utilize apenas uma pessoa. O termo sujeito único se refere ao fato de que a pesquisa é delineada de forma a permitir a comparação rigorosa dos resultados do participante em relação ao seu próprio desempenho. Sendo assim, há pesquisas com esse tipo de delineamento com dois, três ou até uma dezena de participantes. O importante é que o dado de cada participante é analisado separadamente e não tomado em relação à média do grupo.

seria verificada para outras pessoas ou para a mesma pessoa em outras circunstâncias (por exemplo, o mesmo aplicativo para outro conteúdo educacional). Além disso, é possível que a mudança verificada não seja resultado da variável independente, mas de outras variáveis não controladas, como, por exemplo, a proximidade do fim do semestre. Para lidar com essas limitações, foram implementadas variações desse delineamento.

Delineamento A-B-A ou delineamento reverso

Nesse caso, há uma condição inicial de medida de linha de base (A), seguida por uma fase de tratamento (B) e, após esta, retorna-se à condição inicial (A). Com isso, espera-se que, caso as mudanças sejam resultado da intervenção e não de outras variáveis, a situação de reversão (segunda fase A) deverá demonstrar um retorno dos resultados aos níveis iniciais da linha de base (Risley, 2005). Desse modo, identifica-se se as mudanças verificadas na fase B são devidas às variáveis manipuladas, e também pode-se avaliar como o comportamento alvo é alterado quando estas são retiradas de operação.

A limitação desse tipo de delineamento é que ele só se aplica a comportamentos que são reversíveis. Então, por exemplo, o tempo que um aluno passa em um aplicativo pode ser maior ou menor dependendo de características do mesmo. Agora, caso o que esteja sendo avaliado seja, por exemplo, um método de ensino de leitura, não é razoável esperar que a retirada do método implique em que o aluno esqueça o que aprendeu.

Além disso, pode haver considerações éticas envolvidas em retirar o tratamento. Esse seria o caso quando os comportamentos avaliados envolvem ações inadequadas ou indesejáveis. Se uma ferramenta educacional visa diminuir a agressividade do estudante no contexto da sala de aula - e o faz - pode ser considerado antiético retirá-la de operação e, com isso, aumentar seu nível de agressividade.

Delineamento A-B-A-B

Idem ao delineamento anterior, mas com reinserção da fase experimental, a título de replicação dos efeitos da variável independente.

Delineamento de Linha de base múltipla

Nesse delineamento, é realizada uma série de experimentos com delineamento AB para diferentes comportamentos de uma mesma pessoa, diferentes ambientes ou diferentes participantes. O tempo da medida da linha de base costuma ser diferente em cada caso. Assim, verifica-se que efeito da variável de interesse exerce nas diferentes implementações. A variação da linha de base visa demonstrar claramente que os comportamentos alvo se modificaram apenas quando o tratamento se iniciou (BAER; WOLF; RISLEY, 1968) .

Então, ampliando o exemplo anterior, imaginemos que o aplicativo educacional visa ensinar leitura, matemática e ciências e, para isso, têm módulos de ensino independentes. Na fase de linha de base, verifica-se, separadamente, o tempo que o aluno passa em cada um dos módulos até obter-se estabilidade da medida. Na fase seguinte, os melhoramentos propostos são aplicados apenas no módulo de leitura, não afetando os módulos de matemática e ciência. Os três comportamentos continuam sendo mensurados, entretanto, espera-se que, caso o tratamento seja efetivo, apenas o tempo gasto no módulo de leitura apresente alterações. Quando a estabilidade dessa mudança for verificada, o tratamento é introduzido também no módulo de matemática e as medidas são continuadas. Finalmente, o tratamento é inserido no módulo de ciência.

Esse mesmo raciocínio pode ser aplicado utilizando três alunos diferentes ao invés de três módulos do programa. Então, avalia-se a linha de base inicial de todos os alunos e, a partir

da estabilidade, introduz-se a intervenção apenas para um destes; depois para o outro e finalmente para o terceiro. Caso a variável que explica a mudança do comportamento desses participantes seja a intervenção proposta e não qualquer outra coisa (como a aproximação do final de semestre), espera-se que o aumento no tempo de estudo ocorra, para cada um, apenas a partir da introdução da intervenção. Para exemplificar o delineamento de linha de base múltipla entre situações, imaginemos que o aplicativo possa ser usado no computador, no tablet e no celular. Nesse caso, novamente, o tempo de estudo é registrado separadamente para cada um dos aparelhos e as mudanças são inseridas sequencialmente. Se o tratamento for o responsável pela mudança, veremos o tempo de estudo aumentar apenas a partir de sua inserção naquele contexto específico. O delineamento de linha de base múltipla tem muitas vantagens e é considerado um dos meios mais seguros de garantir que os resultados advêm do tratamento proposto. Sendo assim, o modelo apresenta validades internas e externas, além de não exigir reversão e, por isso mesmo, poder ser aplicado a comportamentos que não sejam reversíveis. Suas desvantagens estão relacionadas à aplicabilidade: por exigir estabilidade em múltiplas linhas de base, o trabalho pode demorar bastante tempo. Além disso, requer múltiplas e sistemáticas medidas de comportamentos, situações ou participantes. Entre as variações possíveis para esse modelo, é importante registrar que este pode ser realizado de forma concorrente (simultânea) ou não concorrente (sequencial). Além disso, em casos em que a medida paulatina do comportamento é muito difícil, pode-se recorrer à utilização de sondas de monitoramento, ou seja, medidas pontuais do comportamento de interesse sistematicamente realizadas entre as condições propostas.

Tratamentos alternados

Esse modelo é também denominado delineamento multi-elemento; delineamento de tratamentos múltiplos e delineamento de esquemas múltiplos. Ele permite a comparação do efeito de duas ou mais variáveis independentes ou de dois ou mais tratamentos. Sua realização implica em alternar a aplicação de tratamentos de modo sistemático e contínuo (BARLOW; HAYES, 1979). No caso do exemplo citado, as mudanças no aplicativo educacional, que visam aumentar o tempo total de estudo dos alunos, seriam disponibilizadas dia sim, dia não, por um período determinado de tempo. Posteriormente, o desempenho nos dias com tratamento e sem tratamento seriam comparados.

É importante que o experimento seja planejado cuidadosamente para que a única variação seja o tratamento, porque, caso o aluno faça curso de inglês nas segundas, quartas e sextas, é possível que seu desempenho esteja sendo afetado por essa variável e não pelas mudanças implementadas.

Além disso, usualmente são utilizados modos de "sinalizar" a mudança para os participantes, como instruções, cores etc. Sendo assim, no momento que o aluno acessa o aplicativo, ele sabe que, naquela condição, irá encontrar certas características e, com isso, se sinta mais ou menos motivado a engajar-se no estudo.

As vantagens desse modelo são muitas: ele é de fácil aplicação, não requer estabilidade da linha de base e permite comparação imediata entre tratamentos. Entretanto, é necessário que os tratamentos não sejam muito similares entre si e que um não exerça interferência no outro. Se os dias em que o aplicativo não apresenta os aperfeiçoamentos forem demasiadamente chatos, é possível que ele não volte a interagir com a ferramenta no momento em que o tratamento estiver disponível.

A análise de dados de pesquisas de sujeito único e pequenos grupos

Como os mecanismos de controle desse tipo de pesquisa são muito sistemáticos, é muito frequente que as mudanças verificadas possam ser constatadas apenas por inspeção visual. Essa é uma abordagem para interpretar os resultados de uma pesquisa que implica apenas em inspecionar visualmente o gráfico para averiguar mudanças dentro das fases ou entre estas. Cabe avaliar a intensidade da mudança e a extensão da sobreposição dos resultados. Caso várias das medidas se situem na mesma faixa de variação, a inspeção visual pode ser insuficiente. O mesmo ocorre no caso de existência de tendência de aumento ou diminuição durante a fase de linha de base, uma vez que esta pode interferir nas medidas posteriores. Os resultados de um experimento com delineamento de sujeito único ou pequenos grupos também podem ser avaliados usando estatística, mas, como a quantidade de dados é menor que a de um experimento comum, faz-se necessário utilizar métodos específicos, como Cohen d, para avaliação do tamanho do efeito, Generic Linear Models (GLMs), entre outros. O Quadro 4 apresenta algumas orientações para a descrição de estudos de Sujeito Único ou pequenos grupos. Ela deve ser utilizada em conjunto com a Tabela 1.

Quadro 4 - Orientações Específicas para Relatos de Estudos de Sujeito Único ou Pequenos Grupos

Delineamento	Descreva o delineamento, incluindo <ul style="list-style-type: none">- Tipo de delineamento (por exemplo, retirada-reversão; linha de base múltipla; tratamentos alternados; mudanças de critérios e/ou alguma combinação dos anteriores);- Fases e sequência das fases (determinada a priori ou a partir dos dados obtidos). Se aplicável, apresente os critérios para mudança de fase.
Tipo de delineamento	Mudanças procedimentais <ul style="list-style-type: none">- Descreva qualquer mudança procedimental que ocorreu durante o curso da investigação após o início do estudo. Replicação <ul style="list-style-type: none">- Descreva qualquer replicação planejada; Randomização <ul style="list-style-type: none">- Apresente qualquer randomização que tenha sido usada, descrevendo o método de aleatorização e quais elementos foram aleatorizados.
Análise	Sequência completa <ul style="list-style-type: none">- Relate, para cada participante, a sequência realizada de forma completa, incluindo número de tentativas ou tempo de sessão para cada fase;- Relate os casos de participantes que não completaram a sequência planejada, apresentando as razões para isso. Resultados e estimativas <ul style="list-style-type: none">- Relate resultados para cada participante, incluindo dados brutos para cada desempenho alvo e outros resultados.

Fonte: Adaptado de Appelbaum et al. (2018).

Conclusão

Neste documento apresentamos um guia para apresentação de informações em publicações de pesquisas quantitativas em Informática na Educação. Estas diretrizes são relevantes não apenas para conformidade com a excelência internacional de pesquisas nesta área, mas também para permitir uma melhor apreciação dos resultados. Esta padronização é crucial para realização de revisões sistemáticas e metanálises e revisão de literatura no contexto da produção científica brasileira. Portanto, almejamos contribuir com a disseminação dos resultados da pesquisa brasileira na direção da pesquisa baseada em evidências (BITTENCOURT; ISOTANI, 2018).

Contudo, é importante mencionar que nosso intuito foi iniciar e fomentar a discussão de diretrizes para confecção de artigos científicos em Informática na Educação no Brasil. Assim, acreditamos que colocamos aqui apenas alguns elementos iniciais para esta reflexão e discussão, sem qualquer pretensão de cobrir normatizações em sua totalidade. Este é um campo em formação no Brasil e em constante evolução, de forma que as diretrizes e recomendações aqui apresentadas provavelmente serão melhoradas e reformuladas posteriormente, conforme as necessidades da comunidade científica e órgãos públicos.

Referências Bibliográficas

- APPELBAUM, Mark *et al.* Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, v. 73, n. 1, p. 3–25, jan. 2018. Disponível em: <<http://dx.doi.org/10.1037/amp0000191>>.
- BAER, Donald M.; WOLF, Montrose M.; RISLEY, Todd R. Some current dimensions of applied behavior analysis 1. *Journal of Applied Behavior Analysis*, v. 1, n. 1, p. 91-97, 1968.
- BARLOW, David H.; HAYES, Steven C. Alternating treatments design: One strategy for comparing the effects of two treatments in a single subject. *Journal of Applied Behavior Analysis*, v. 12, n. 2, p. 199-210, 1979.
- BITTENCOURT, Ig Ibert; ISOTANI, Seiji. Informática na Educação baseada em Evidências: um Manifesto. *Revista Brasileira de Informática na Educação*, v. 26, n. 3, p. 108, 13 set. 2018. Disponível em: <<http://www.br-ie.org/pub/index.php/rbie/article/view/7826>>.
- COOPER, J.O; HERON, T.E; HEWARD, W.L. *Applied behavior analysis*. New York: Macmillan, 1987.
- DERMEVAL, D.; COELHO, J. A. P. de M.; BITTENCOURT, I. I. Mapeamento Sistemático e Revisão Sistemática da Literatura em Informática na Educação. In: JAQUES, Patrícia Augustin; PIMENTEL, Mariano; SIQUEIRA, Sean; BITTENCOURT, Ig. (Org.) *Metodologia de Pesquisa em*

- Informática na Educação: Abordagem Quantitativa de Pesquisa. Porto Alegre: SBC, 2019. (Série Metodologia de Pesquisa em Informática na Educação, v. 2). Disponível em: <<https://metodologia.ceie-br.org/livro-2>>.
- JAQUES, Patrícia Augustin; PIMENTEL, Mariano; SIQUEIRA, Sean; BITTENCOURT, Ig. (Org.) Série Metodologia de Pesquisa em Informática na Educação. Porto Alegre: SBC, 2019. Disponível em: <<https://metodologia.ceie-br.org/>>.
- GAZZANIGA, M; HEATHERTON, T. Ciência Psicológica. Porto Alegre, RS: Artmed Editora, 2005.
- GREER, Jim E; MARK, Mary A. Evaluation Methods for Intelligent Tutoring Systems Revisited. *International Journal of Artificial Intelligence in Education*, v. 26, n. 1, p. 387–392, 23 mar. 2016. Disponível em: <<http://link.springer.com/10.1007/s40593-015-0043-2>>.
- KARUMBIAIAH, Shamyia *et al.* The Implications of a Subtle Difference in the Calculation of Affect Dynamics. *Proceedings of the 26th International Conference on Computers in Education*, n. November, 2018.
- KUMAR, Ranjit. Research methodology: A step-by-step guide for beginners. Sage Publications Limited, 2019.
- LEVITT, Heidi M *et al.* Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology. *American Psychologist*, v. 73, n. 1, p. 26–46, 2018. Disponível em: <<http://dx.doi.org/10.1037/amp0000151>>.
- MARK, Mary A; GREER, Jim E. Evaluation methodologies for intelligent tutoring systems. *Journal of Artificial Intelligence and Education*, v. 4, n. 2/3, p. 129–153, 1993. Disponível em: <<http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Evaluation+Methodologies+for+Intelligent+Tutoring+Systems#0>>. Acesso em: 5 ago. 2014.
- NICHOLS, Thomas E. et al. Best practices in data analysis and sharing in neuroimaging using MRI. *Nature neuroscience*, v. 20, n. 3, p. 299, 2017.
- RISLEY, Todd. Montrose M. Wolf (1935–2004). *Journal of Applied Behavior Analysis*, v. 38, n. 2, p. 279-287, 2005.
- VOHRA, S.; SHAMSEER, L.; SAMPSON, M.; BUKUTU, C.; SCHMID, C. H.; TATE, R., . . . CENT group. CONSORT extension for reporting N-of-1 trials (CENT) 2015 statement. *British Medical Journal*, v. 350, p. h1738, 2015.