

# Mapeamento Sistemático sobre o uso do Learning Analytics na Aprendizagem Infantojuvenil com ênfase em Jogos

Alexandre Mendonça Fava<sup>1</sup>

<sup>1</sup>Universidade do Estado de Santa Catarina - UDESC  
Programa de Pós-graduação em Computação Aplicada - PPGCA  
Joinville - SC - Brasil CEP: 89.219-710

{alexandre.fava@hotmail.com}

**Resumo.** *A educação a distância é uma modalidade educativa em expansão. O processo de educação a distância gera uma quantidade superior de informações em relação ao ensino tradicional. O Learning Analytics é uma área de pesquisa relacionada à coleta e análise de dados dos alunos para o aprimoramento do aprendizado escolar. A área de Learning Analytics compreende inúmeras técnicas, objetivos e grupos. Nesse sentido, o presente artigo apresenta um mapeamento sistemático da literatura na área de Learning Analytics com ênfase em jogos no aprendizado de menores de idade (18 anos inclusive). O resultado final, é uma compilação acerca dos procedimentos, campos e instrumentos de maior relevância na área em questão.*

## 1. Introdução

O uso das mídias digitais e dispositivos conectados a grande rede de computadores tem crescido desde o início da internet comercial no Brasil (Comitê Gestor da Internet no Brasil, 2019). Em 2020, a pandemia de COVID-19 foi responsável por um súbito crescimento nas plataformas de ensino a distância, tais como Google Classroom, Microsoft Teams e Blackboard Learn (MONTEIRO, 2020).

As plataformas de educação a distância demonstraram-se como uma alternativa viável na contenção e na propagação do vírus que causa a COVID-19 (SARS-CoV-2). A diminuição da circulação e do contato entre as crianças e jovens, no ambiente escolar, é uma atitude capaz de reduzir a taxa de contágio do vírus (FERGUSON et al., 2020).

Diferentemente das salas de aula tradicionais, as plataformas de ensino a distância permitem uma coleta mais automatizada e individualizada de algumas informações. Por exemplo: a verificação da presença dos alunos em uma determinada aula pode ser conferida nos arquivos de *log* das plataformas, dispensando a necessidade de uma ‘chamada’.

As informações coletadas não se limitam apenas às informações nas quais os professores já poderiam coletar nas aulas tradicionais. As salas de aula virtuais permitem não apenas uma coleta mais automatizada, mas também uma coleta mais abrangente de informações, o que pode acabar dificultando o acompanhamento do professor em analisar todos os dados estudantis gerados pela plataforma (SOUZA; WIVES; PERRY, 2019). No mais, o sistema educacional atual não se limita apenas nas aulas com suas diapositivas e exercícios. O processo de educação abrange atualmente vários recursos e tecnologias, como por exemplo a área de jogos sérios (FIOCRUZ, 2020).

Em vias de facilitar o processo de análise e compreensão dos dados estudantis, técnicas de Learning Analytics (LA) surgiram como um suporte para as tarefas de análise de aprendizagem virtual (RUIPÉREZ-VALIENTE et al., 2015). O processo de LA busca coletar, medir, analisar e relatar os dados e seus contextos com objetivo de otimizar o aprendizado e o ambiente em que este ocorre (MOISSA; GASPARINI; KEMCZINSKI, 2015). Percebe-se neste processo que os alunos obtêm abordagens customizadas para auxiliá-los a aprender e base de informação em um método que se encaixe às suas necessidades e não à classe toda.

Apresentada a emergência das plataformas de educação digitais, e a importância do processo de LA para facilitar a ensino-aprendizagem, a atual pesquisa se objetiva a apresentar um panorama da área em questão com enfoque em jogos. O restante do trabalho está organizado da seguinte forma: na seção 2 são elencados alguns trabalhos relacionados, na seção 3 o mapeamento e seus processos são especificados, na seção 4 os principais achados são compilados e na seção 5 são apresentadas as conclusões desta pesquisa.

## **2. Trabalhos Relacionados**

Moissa, Gasparini e Kemczinski (2014) realizaram um mapeamento sistemático na área de Learning Analytics. Ao total foram analisados 116 trabalhos compreendidos entre os anos de 2011 e 2014. A análise dos trabalhos revelou que a coleta de dados mais utilizada na área de Learning Analytics se dá por meio de dados navegacionais, tais como tempo *online* nas ferramentas e frequência de acesso. O histórico escolar das crianças também foi outra variável bem presente no mapeamento. Entre os achados, está a conclusão de que dados coletados de mídias sociais também podem auxiliar na melhor compreensão do comportamento dos alunos e auxiliar no gerenciamento da aprendizagem colaborativa.

Doko e Bexheti (2018) coletaram 122 artigos nas temáticas de Mineração de Dados na educação e Learning Analytics. Um dos resultados diretos do artigo, foi a observação de um considerável crescimento da área entre os anos de 2010 e 2017. Entre os maiores achados, o mapeamento constatou uma crescente preferência por *video segments* nas pesquisas.

Damasceno et al. (2018) analisaram vários artigos em língua inglesa e portuguesa publicados após 2009, ao final 109 artigos foram analisados. Como resultado, foi constatado comparativamente que o objetivo da maioria dos artigos era o de identificar padrões de comportamento e trajetória dos alunos, sendo a matéria mais vista nesse contexto, a Informática. Matérias mais tradicionais como Matemática e Química não apresentaram larga presença no mapeamento realizado.

Os trabalhos relacionados demonstram convergência no que diz respeito ao volume de artigos analisados. Do mesmo modo, os mapeamentos aqui elencados não relataram muitos jogos na área. Salienta-se que todos os trabalhos apresentados firmaram suas buscas nos buscadores acadêmicos da ACM e da IEEE Xplore. Doko e Bexheti (2018) ainda afirmam que a maior parte dos trabalhos colhidos em sua pesquisa é oriunda destes buscadores. Em adendo, revisões sistemáticas da literatura realizadas especificamente na área de jogos sérios relatam também uma quantidade significativa de artigos encontrados nas bases da ACM, Science Direct (Elsevier) e IEEE Xplore (CALVO-MORATA et al., 2020). Nesse contexto, cita-se a revisão da literatura realizada por Calvo-Morata et al. (2020), a qual lista 31 artigos com o envolvimento de crianças. Por tal razão, se espera que a quantidade de artigos elencadas neste mapeamento seja próxima deste valor.

### 3. Mapeamento Sistemático

Mapeamentos sistemáticos são projetados para dar um panorama geral de uma determinada área, envolvendo uma busca na literatura para descobrir os estudos primários que foram publicados em um tema de pesquisa (CHEN, 2017). O mapeamento se diferencia da revisão sistemática, a qual avalia a força das evidências. O processo se divide em: definir uma pergunta de pesquisa, definir palavras-chave, selecionar os estudos, extrair os dados, analisar, classificar e validar (KHAN et al., 2003).

O presente trabalho realiza um mapeamento sistemático na área de Learning Analytics para o ensino infantil com foco na temática de jogos. A área de jogos foi eleita para o presente mapeamento, uma vez constatada sua fraca presença no estudo dos trabalhos relacionados (Seção 2). Deste modo, busca-se ampliar a visão geral da área de LA no campo dos jogos para a educação infantojuvenil.

Os passos necessários para a execução de um mapeamento sistemático na área de LA para a educação infantojuvenil na temática de jogos é descrito mais detalhadamente em suas subseções: a subseção 3.1 apresenta as questões de pesquisa, a subseção 3.2 mostra a frase de busca, a subseção 3.3 define os critérios objetivo para inclusão e exclusão de artigos, assim como o critério de parado do presente mapeamento.

#### 3.1. Questões de Pesquisa

O primeiro passo para a execução de qualquer mapeamento sistemático é a definição das questões de pesquisa. Essas questões de pesquisa são responsáveis por ditar o que se espera conseguir do mapeamento sistemático. As questões que norteiam o atual mapeamento sistemático são:

1. Quais os tipos de dados coletados?
2. Quais são os tamanhos das amostras nos estudos?
3. Quais são as disciplinas mais presentes nos estudos?
4. Quais métodos são utilizados para analisar os alunos?
5. Quais ferramentas são utilizadas para analisar as informações dos estudantes?

A **primeira questão** busca identificar as variáveis que mais se destacam no âmbito investigado. A primeira questão, se mescla um pouco com a **quinta questão**, a qual versa sobre as tecnologias utilizadas para coletar as variáveis mensuradas pelos estudos. Do mesmo modo, a **quarta questão** assume uma posição equivalente visando identificar as estratégias utilizadas para se trabalhar com os dados colhidos. Por fim, a **segunda questão** de pesquisa e a **terceira questão** de pesquisa assumem características mais triviais, as quais possibilitam a identificação não apenas da extensão dos grupos avaliados nas pesquisas, mas também das matérias escolares mais presentes neste âmbito.

#### 3.2. Frase de Busca

Após a definição das questões de pesquisa é necessário estabelecer as bases de dados acadêmicas que serão utilizadas. Buchinger, Cavalcanti e Hounsell (2014) realizaram uma análise quantitativa com 40 Mecanismos de Busca Acadêmicas. Para a execução de uma pesquisa científica é recomendado a utilização de pelo menos três fontes de dados distintas. Para o presente mapeamento, quatro bases foram utilizadas: ACM, DBLP, IEEE Xplore e Science Direct (Elsevier). As bases elencadas, além de serem bons buscadores de acordo com a pesquisa de Buchinger, Cavalcanti e Hounsell (2014), também estão em conformidade com as relatadas nos trabalhos relacionados (Seção 2).

A realização de uma pesquisa nos mecanismos de busca acadêmicas, pode ser realizada com maior precisão com o auxílio de alguns recursos de busca disponíveis como caractere coringa e operadores booleanos. A busca deve ser realizada nas palavras-chave, título e resumo do artigo preferencialmente. Salienta-se que algumas ferramentas de busca estão limitadas a um conjunto máximo de oito operadores booleanos (como o Science Direct). O presente mapeamento formulou uma frase de busca simples com o intuito de respeitar as limitações de todos os mecanismos de busca utilizados:

“Learning Analytics” AND children AND game

A frase de busca utilizada apresenta de maneira simples e em língua inglesa os três termos de maior relevância para o presente mapeamento. Notoriamente, sabe-se que existem outras palavras (sinônimos) para os termos elencados, contudo constatou-se em um primeiro momento conflito na formatação do operador OR entre alguns mecanismos, deste modo optou-se em utilizar apenas o operador lógico AND, o qual permitiu a construção de uma frase de busca invariável entre os mecanismos de busca acadêmica utilizados. Além disso, os termos elencados não fogem do escopo, abrangendo área, técnica e grupo, como recomendado por Wazlawick (2009). O total de publicações retornadas é apresentado detalhadamente na Tabela 1.

**Tabela 1. Quantidade de publicações retornadas por Mecanismo de Busca**

| <b>Mecanismo de Busca</b> | <b>Quantidade</b> |
|---------------------------|-------------------|
| ACM                       | 130               |
| DBLP                      | 1                 |
| IEEE Xplore               | 2                 |
| Science Direct (Elsevier) | 44                |
| <b>Total</b>              | <b>177</b>        |

A busca realizada nos mecanismos elencados retornou um total de 177 resultados. Informa-se nesse sentido que a busca foi realizada no mês de agosto de 2020. Buscas realizadas em uma janela de tempo superior a este período tendem a retornar quantidades cada vez mais diferentes no número de resultados na medida que a data buscada for se afastando da data de busca realizada pela presente pesquisa. Afim de trazer maior transparência aos procedimentos deste mapeamento, enfatiza-se que as páginas dos resultados da busca foram gravadas em PDF e encontram-se disponíveis a qualquer um que queira requisitá-las para conferir os vereditos desta pesquisa.

### 3.3. Critérios Objetivos

A execução da busca dos artigos e demais pesquisas nas bases de dados acadêmicas pode retornar trabalhos indesejados para o mapeamento sistemático. Deste modo surgem os critérios de inclusão e exclusão com o objetivo de ajudar a descartar os artigos que, embora contivessem as palavras-chave definidas na frase de busca, não contribuíam para responder as questões de pesquisa. Como critérios de inclusão estabelece-se:

- Pesquisas publicadas após 2010, inclusive.
- Pesquisas de acesso livre e gratuito.
- Pesquisas envolvendo menores.
- Pesquisas utilizando jogos.
- Pesquisas primárias.

Informa-se que o critério do **ano da publicação** foi o único já englobado no momento da busca, devido a existência de recursos de seleção por período nos mecanismos utilizados. Salienta-se então que todos os 177 artigos retornados pelas bases acadêmicas são de uma data superior ao ano de 2010, inclusive. Este critério vem com o intuito de excluir pesquisas mais antigas, sendo o objetivo do presente mapeamento uma apresentação mais atual do estado da arte da área e também das tecnologias mais recentes utilizadas.

A **disponibilidade** das publicações é um critério importante, o qual permite que os resultados apresentados por esse mapeamento possam ser mais facilmente acessados. A participação de **menores** (indivíduos de até 18 anos) é outro critério necessário para o presente mapeamento, uma vez que existe a possibilidade de pesquisas conterem o termo ‘*children*’ em algum dos campos buscados, sem que estas tenham de fato a participação ou o envolvimento de crianças ou jovens, o mesmo se faz válido no que diz respeito a utilização de **jogos** na pesquisa.

O último critério de inclusão, diz respeito a contribuição do artigo. Deste modo, o presente mapeamento valoriza as **pesquisas primárias**, as quais apresentam novos resultados e experimentos. Todos os critérios de inclusão devem ser verdadeiros para que um artigo seja considerado viável para ser incluído no mapeamento sistemático. Contudo, basta um critério de exclusão ser constatado para uma pesquisa ser descartada. Como critérios de exclusão elencam-se:

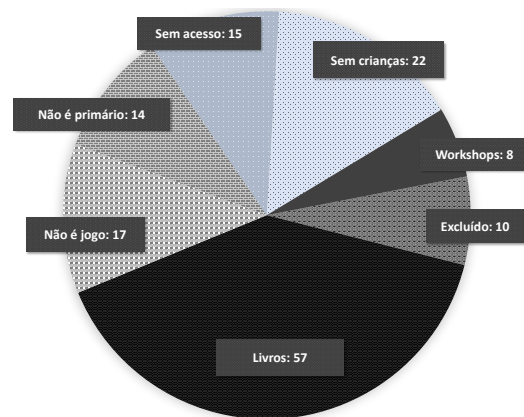
- Pesquisas duplicadas.
- Pesquisas não escritas em inglês.
- Pesquisas sem os termos da frase de busca no título, resumo ou palavra-chave.

A preferência de pesquisas em **língua inglesa**, se deve pelo fato deste idioma apresentar a maior quantidade de publicações acadêmicas do ocidente, em comparação a qualquer outro idioma, o que amplia a quantidade de pesquisas retornadas. A exclusão de **pesquisas repetidas** visa trazer uma melhor acurácia aos numerativos apresentados pelo presente mapeamento, já os campos utilizados para a busca visam trazer maior agilidade para o processo de seleção e qualificação dos artigos, uma vez que todos os processos do presente mapeamento sistemático foram executados por somente um indivíduo. Em adendo, salienta-se que todo o protocolo do mapeamento, tanto quanto seus resultados foram validados pelos orientadores do presente pesquisador.

As 177 publicações retornadas tiveram seus títulos, resumos, introduções e conclusões (quando existentes) lidas de modo a incluir ou excluir uma determinada pesquisa do mapeamento. Nenhum critério de priorização por *Qualis*, veículo de publicação ou quantidade de referências foi aplicado, garantindo a leitura de todas as 177 pesquisas.

Durante o processo de leitura, algumas publicações de revisão da literatura na área foram descobertas (CALVO-MORATA et al., 2020; ALONSO-FERNÁNDEZ et al., 2019). Salienta-se que no decorrer do processo, não houve nenhuma inclusão Ad-hoc e também não foi realizada nenhuma pesquisa ao entorno, tudo isso para garantir e preservar o rigor do processo sistemático do mapeamento. Deste modo, o critério de parada do presente mapeamento foi atingido com a finalização da leitura dos 177 trabalhos, restando 29 artigos ao final. O resultado compilado dos numerativos de exclusão é apresentado de maneira geral na Figura 1. A Tabela 2 apresenta a relação das pesquisas remanescentes por mecanismo de busca.

| <b>Tabela 2. Publicações qualificadas</b> |                     |
|---|---------------------|
| <b>Mecanismo de Busca</b>                 | <b>Qualificados</b> |
| ACM                                       | 20                  |
| DBLP                                      | 1                   |
| IEEE Xplore                               | 1                   |
| Science Direct (Elsevier)                 | 7                   |
| <b>Total</b>                              | <b>29</b>           |



**Figura 1. Publicações excluídas**

A Figura 1 apresenta um gráfico circular com os quantitativos das publicações excluídas após a etapa de leitura do protocolo de mapeamento. Cerca de 33% (57) das publicações retornadas eram livros, 13% (22) não apresentavam o envolvimento de crianças na pesquisa, 10% (17) não utilizavam jogos, 9% (15) não eram acessíveis, 8% (14) não eram pesquisas primárias, 4% (8) eram Workshops e 6% (10) foram excluídos por não conterem os termos buscados nos campos designados. Ao final, 17% (29) foram validados pelo processo de leitura.

Todas as etapas do presente mapeamento foram executadas sem conflitos de interesse. A agência financiadora da presente pesquisa (CAPES) dá livre liberdade de expressão, manifestação e pesquisa para seus bolsistas. Contudo, salienta-se que o atual mapeamento (executado com protocolo bem definido e seguido rigorosamente) não é imune a falhas. Eventuais equívocos ou descuidos podem ter ocorrido durante o processo, no que diz respeito a validade interna e externa do mapeamento.

Os problemas de validade externa podem ter ocorrido devido à falta de padronização em alguns termos utilizados nas frases de busca. Sabe-se que o termo mais academicamente aceito para jogos no contexto educacional é a terminologia “jogo sério”, cunhada em 2002 (DJAOUTI et al., 2011). Entretanto, pesquisas mais recentes continuam a utilizar outros termos, uma vez que não houve a padronização da terminologia. Afirma-se que a palavra “jogo” encontra-se presente também nas demais terminologias, por tal razão acredita-se que os problemas de integridade externa tenham sido minimizados nesse sentido.

Os problemas de integridade interna podem ser inúmeros. No contexto desta pesquisa o problema de maior destaque nesse sentido, vem no momento da apresentação dos grupos avaliados em determinados trabalhos. Grande parte dos trabalhos definiam claramente um limite inferior e superior da idade dos participantes da pesquisa, o que ajuda na condensação dos dados para o presente mapeamento. Todavia, um conjunto de trabalhos apresentava a idade dos participantes de maneira indireta, mencionando suas turmas ou classes de aula por exemplo. A falta de clareza neste sentido, foi um dos fatores que contribuíram para possíveis problemas de integridade interna no mapeamento realizado.

Todo o processo de seleção, leitura, coleta e análise dos trabalhos foi realizado no período de um mês. A execução do protocolo de mapeamento e a escrita do presente artigo foi realizada por um bolsista da CAPES. Salienta-se novamente que o protocolo de mapeamento foi validado pela orientadora e coorientador do presente autor.

## 4. Resultados e Análise

O último passo para o mapeamento sistemático é o de analisar o conteúdo dos artigos restantes, encontrando relações, achados importantes e respondendo as questões de pesquisa formuladas. Durante esse processo artigos podem ser descartados, mas não incluídos. A leitura completa dos 29 artigos restantes foi realizada, nesse processo nenhum artigo foi excluído. A separação dos artigos por veículo de publicação é mostrada na Figura 2. Já a dispersão temporal de todos os artigos é apresentada mais detalhadamente na Figura 3.

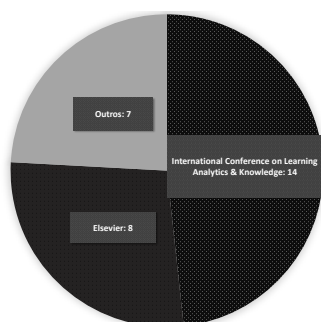


Figura 2. Separação dos artigos por local de publicação

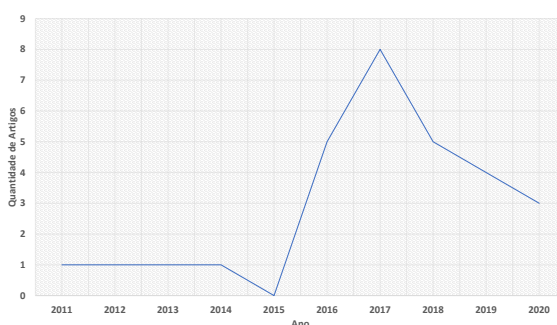


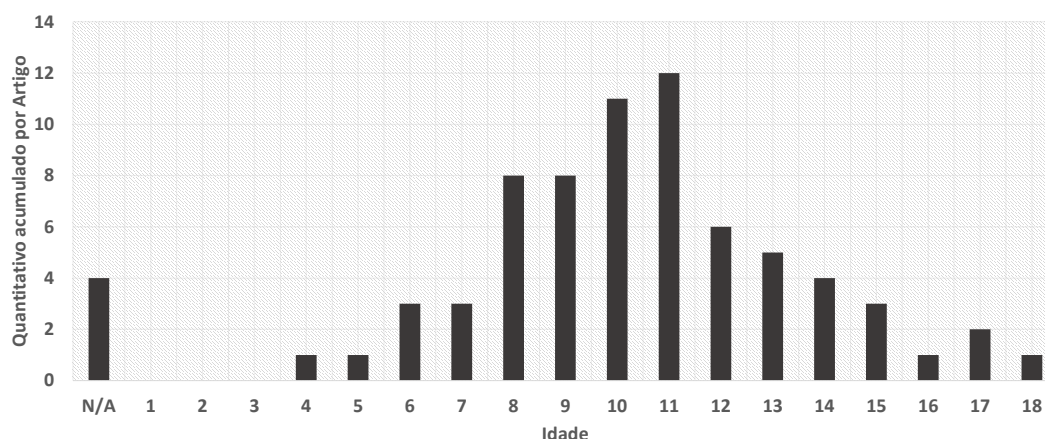
Figura 3. Quantidade de artigos por ano de publicação

A Figura 3 demonstra um considerável crescimento de estudos que relacionam jogos e Learning Analytics com o envolvimento de menores após 2015. Reitera-se nesse sentido que os dados para o ano de 2020 vão somente até o mês de agosto, o que resulta um quantitativo incompleto para o ano em questão. A Figura 2 apresenta os locais onde as pesquisas foram publicadas. A conferência de maior destaque nesse contexto é a *International Learning Analytics and Knowledge Conference*. Outras conferências também marcaram presença no presente mapeamento, como a *International Conference on Interaction Design and Children* e a *Conference on Innovation and Technology in Computer Science Education*.

A quantidade de participantes envolvidos nos estudos apresentou grande flutuação. Os maiores numerativos nesse sentido foram de pesquisas que se utilizaram de bases de dados já existentes. A maior base utilizou dados de mais de 300.000 crianças, tais dados foram extraídos do sistema Todo Math<sup>1</sup> (KIM et al., 2018). O menor numerativo encontrado durante o mapeamento foi de um **estudo de caso** que acompanhou por cinco dias uma única menina de dez anos de idade (PANTIC; FIELDS; QUIRKE, 2016).

Mais de 50% das pesquisas avaliaram grupos de dez até sessenta indivíduos. Cerca de 18% avaliaram grupos de 100 até 500 indivíduos. Pouco mais de 10% avaliaram grupos menores de 10 indivíduos. Os 12% restantes se enquadram em outros numerativos. Metade dos trabalhos informaram as dimensões de seus grupos por gênero. Com a exceção de quatro artigos, todos os demais relataram as faixas etárias dos grupos avaliados. O gráfico de barras da Figura 4 mostra a dispersão por idade dos estudos.

<sup>1</sup>Todo Math é um aplicativo de aprendizagem baseado em dispositivos móveis que contém conceitos matemáticos básicos para crianças nas fases iniciais do ensino fundamental.

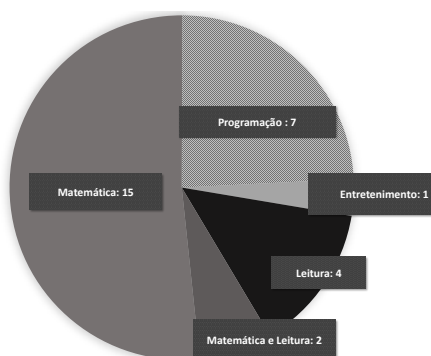


**Figura 4. Distribuição da idade dos participantes dos artigos**

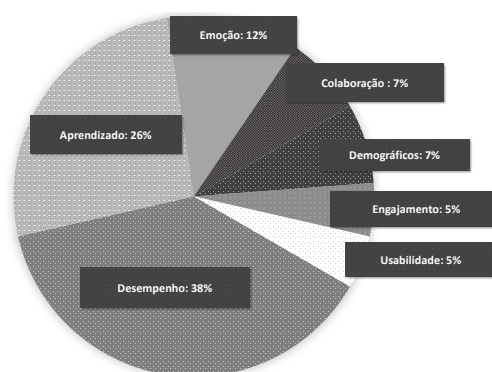
A Figura 4 apresenta a variação de idade encontrada ao longo do presente mapeamento. A sigla **N/A** (Não se Aplica) representa os artigos nos quais não foi possível mensurar a faixa etária dos envolvidos no estudo. Parte dos trabalhos relatavam somente o ano letivo dos participantes, onde a idade dos participantes só foi possível mensurar de maneira indireta. Os trabalhos relatados em **N/A**; ou não apresentavam nenhum informativo do qual a informação da idade dos participantes poderia ser inferida; ou relatavam dados imprecisos (contudo, foi possível inferir que tratavam-se de menores de 18 anos).

A distribuição etária da Figura 4 mostra uma distribuição normal, onde a faixa etária de maior destaque é a faixa dos onze anos de idade. Salienta-se que os intervalos etários convergiam entre algumas pesquisas, por tal razão o numerativo total da distribuição etária é superior a quantidade total de artigos avaliados. As exceções ocorrem nos artigos que não mencionaram um intervalo da idade, mas sim, a média de idade dos participantes. É importante destacar que os baixos numerativos nas extremidades do gráfico podem ter ocorrido em virtude dos termos utilizados na frase de busca, sendo o termo *toddler* o mais apropriado para descrever as crianças de zero até três anos de idade e o termos *teenager* ou *adolescent* para as de treze até dezenove anos.

A distribuição etária reflete nas temáticas dos jogos presentes nos estudos e nas variáveis coletadas, tais quantitativos podem ser observados mais detalhadamente na Figura 5 e na Figura 6.



**Figura 5. Artigos por assunto**



**Figura 6. Percentual das variáveis mensuradas nos artigos**



A Figura 5 mostra as principais temáticas dos jogos no mapeamento realizado. Grande maioria dos jogos eram de propósito didático, a exceção está na pesquisa de Tsuei et al. (2019) que realizou uma avaliação de usabilidade de um jogo de entretenimento VR (Realidade Virtual). A Figura 6 condensa as variáveis de maior expressão mensuradas pelos estudos. Salienta-se que grande parte das pesquisas relacionavam duas ou mais variáveis, sendo a relação de maior destaque a variável de **desempenho** das crianças no jogo com a variável de **aprendizado** dos conteúdos no âmbito escolar.

Para coletar e mensurar as variáveis, inúmeras técnicas foram utilizadas pelos estudos. A grande maioria se utilizaram de *log de dados* (14 artigos). Coletas de áudio e vídeo também estiveram presentes em alguns estudos (8 artigos). A utilização de testes também é relatada em algumas pesquisas (20 artigos), porém uma quantidade menor de pesquisas relataram a aplicação tanto de pré, quanto de pós-testes (10 artigos), além disso, um numerativo reduzido de artigos informaram se utilizar de um grupo controle para isolar variáveis (6 artigos). Da mesma forma, poucas pesquisas relataram a utilização de *hardwares* ou *softwares* mais avançados para a coleta ou análise das variáveis (4 artigos), tais como: Eye-Tracking, Wristband, FaceReader e Kinect.

## 5. Conclusão

Neste trabalho foi realizado um mapeamento sistemático conduzido nas bases acadêmicas ACM, DBLP, IEEE Xplore e Science Direct (Elsevier). Ao total 177 publicações foram retornadas; após a filtragem das publicações, 29 artigos remanesceram. Os artigos restantes permitiram identificar quais são os dados mais almejados pelas pesquisas na área da educação infantil e juvenil com auxílio de jogos. Tais dados, precisam ser coletados e devidamente analisados, deste modo surge o Learning Analytics que tem como objetivo melhorar o processo de ensino-aprendizagem por meio da análise de dados estudantis. Dentre as informações mais almejadas pelas pesquisas estão: a coleta de relatório de desempenho, aprendizado e emoções. Isso permite responder a primeira questão de pesquisa do presente mapeamento.

A coleta das variáveis pode ser realizada de inúmeras maneiras, nos artigos elencados as formas de coleta mais presentes nos estudos são: coleta por *log de dados*, materiais audiovisuais e testes/provas. Isso permite responder a quinta questão de pesquisa deste mapeamento. Informa-se que tal achado está em consonância com a pesquisa de Moissa, Gasparini e Kemczinski (2014), a qual relatada também em seu mapeamento grande presença de pesquisas que se utilizam de dados navegacionais (*log de dados*). Também há consonância com a pesquisa de Doko e Bexheti (2018) que relatou grande presença na utilização de *videos segments* nas pesquisas.

Após a coleta dos dados, uma análise precisa ser realizada. Nas pesquisas elencadas a análise dos dados ou era realizada de maneira subjetiva, ou com o auxílio de um algoritmo não documentado. Poucos artigos elencados se utilizaram de formas de análise de dados mais reconhecidas na academia, tais como: Eye-Tracking, FaceReader e a Taxonomia de Bloom. Isso auxilia a responder a quarta questão de pesquisa.

No decorrer desta pesquisa observou-se que as disciplinas escolares de maior presença nos estudos são, em ordem decrescente: Matemática, Programação e Leitura (Língua Inglesa). Tal informação responde a terceira questão de pesquisa deste mapeamento, porém contrasta um pouco com o mapeamento realizado por Damasceno et al.

(2018), o qual relata mais chamados nas matérias de informática e programação do que matérias como a matemática. Por fim, para responder a segunda questão de pesquisa do atual mapeamento se buscou o quantitativo numérico de indivíduos nos estudos, em resposta, mais da metade dos trabalhos elencados, continuam grupos de 10 até 60 indivíduos.

A maioria das pesquisas nessa área buscam identificar relações entre o desempenho infantil durante o jogar do jogo e seu aprendizado escolar. Salieta-se que as pesquisas colhidas pelo presente mapeamento apresentam resultados promissores nesse sentido, apontando relação de sapiência entre jogos e alguns conceitos disciplinares. A disciplina mais relatada nos estudos foi a matemática. Tais resultados contrastam com a baixa expressividade de publicações na área em questão. Os resultados das pesquisas apresentam conclusões promissoras, contudo há baixa expressividade de publicações na área, o que implica em um campo inteiro ainda para ser devidamente explorado e estudado por pesquisas futuras.

## **Agradecimentos**

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

## **Referências**

- ALONSO-FERNÁNDEZ, C. et al. Applications of data science to game learning analytics data: A systematic literature review. *Computers & Education*, Elsevier, v. 141, p. 103612, 2019.
- BUCHINGER, D.; CAVALCANTI, G. A. de S.; HOUNSELL, M. da S. Mecanismos de busca acadêmica: uma análise quantitativa. *Revista Brasileira de Computação Aplicada*, v. 6, n. 1, p. 108–120, 2014.
- CALVO-MORATA, A. et al. Serious games to prevent and detect bullying and cyberbullying: A systematic serious games and literature review. *Computers & Education*, Elsevier, p. 103958, 2020.
- CHEN, C. Science mapping: a systematic review of the literature. *Journal of Data and Information Science*, Sciendo, v. 2, n. 2, p. 1–40, 2017.
- Comitê Gestor da Internet no Brasil. Pesquisa sobre o uso das tecnologias de informação e comunicação nos domicílios brasileiros: 2018. *São Paulo: Comitê Gestor da Internet no Brasil*, 2019.
- DAMASCENO, A. de B. et al. Entrevistas e mapeamento sistemático sobre análise de logs de interação de alunos em ambientes virtuais de aprendizagem. *Monografias em Ciência da Computação*, v. 1200, n. 08, 2018.
- DJAOUTI, D. et al. Origins of serious games. In: *Serious games and edutainment applications*. [S.l.]: Springer, 2011. p. 25–43.
- DOKO, E.; BEXHETI, L. A. A systematic mapping study of educational technologies based on educational data mining and learning analytics. In: *IEEE. 2018 7th Mediterranean Conference on Embedded Computing (MECO)*. [S.l.], 2018. p. 1–4.

FERGUSON, N. et al. Report 9: Impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand. *Imperial College London*, v. 10, p. 77482, 2020.

FIOCRUZ. Ensino a distância na educação básica frente à pandemia da covid-19. 2020.

KHAN, K. S. et al. Five steps to conducting a systematic review. *Journal of the royal society of medicine*, SAGE Publications Sage UK: London, England, v. 96, n. 3, p. 118–121, 2003.

KIM, B. et al. "i'll do it!"examining the relationship between locus of control and math game retention for preschoolers. In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. [S.l.: s.n.], 2018. p. 290–294.

MOISSA, B.; GASPARINI, I.; KEMCZINSKI, A. Learning analytics: um mapeamento sistemático. *Nuevas Ideas en Informática Educativa TISE*, v. 2014, p. 283–290, 2014.

MOISSA, B.; GASPARINI, I.; KEMCZINSKI, A. Educational data mining versus learning analytics: estamos reinventando a roda? um mapeamento sistemático. In: *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*. [S.l.: s.n.], 2015. v. 26, n. 1, p. 1167.

MONTEIRO, S. da S. Inventar educação escolar no brasil em tempos da covid-19. *Revista Augustus*, v. 25, n. 51, p. 237–254, 2020.

PANTIC, K.; FIELDS, D. A.; QUIRKE, L. Studying situated learning in a constructionist programming camp: A multimethod microgenetic analysis of one girl's learning pathway. In: *Proceedings of the The 15th International Conference on Interaction Design and Children*. [S.l.: s.n.], 2016. p. 428–439.

RUIPÉREZ-VALIENTE, J. A. et al. Alas-ka: A learning analytics extension for better understanding the learning process in the khan academy platform. *Computers in Human Behavior*, Elsevier, v. 47, p. 139–148, 2015.

SOUZA, N. S. de; WIVES, L. K.; PERRY, G. T. Tendências de pesquisas que utilizam learning analytics em moocs: um mapeamento sistemático. *RENOTE-Revista Novas Tecnologias na Educação*, v. 17, n. 1, p. 82–92, 2019.

TSUEI, M. et al. Preliminary evaluation of the usability of a virtual reality game for mudslide education for children. In: *25th ACM Symposium on Virtual Reality Software and Technology*. [S.l.: s.n.], 2019. p. 1–2.

WAZLAWICK, R. S. *Metodologia de pesquisa para ciência da computação*. [S.l.]: Elsevier, 2009.