

## Análise dos sistemas para buscas sistemáticas na área de ensino no Brasil

Flávia Linhalis\*  
Bruno Xavier Duarte†

NIED – Relatório Técnico Nº 1  
2025

### Resumo

As revisões sistemáticas, também referidas como revisões de literatura ou revisões bibliográficas, requerem a estratégia de pesquisa correta, usando bases de dados apropriadas e sistemas de busca eficientes. Na área de Ensino, existe uma carência por publicações sobre quais sistemas de busca são adequados para buscas sistemáticas. Este estudo mostra uma avaliação com análise das qualidades de busca e recuperação de 10 sistemas de busca acadêmica amplamente utilizados para realizar buscas sistemáticas por pesquisadores brasileiros, considerando a área de Ensino. A metodologia para análise dos sistemas de busca inclui 27 critérios divididos nas categorias: abrangência, busca, resultados da busca e reprodutibilidade da busca. Os resultados mostram a análise do Portal de Periódicos da CAPES, SciELO.br, Pubmed, Scopus, Web of Science (WoS), BDTD, LILACS, Banco de Teses e Dissertações da CAPES, Google Acadêmico e ERIC. Os resultados mostram que, embora alguns sistemas possam ser recomendados quase sem limitação, outros falharam em testes importantes, que limitam sua adequação para buscas sistemáticas. Os resultados indicam que três dos dez sistemas avaliados não são adequados como fontes primárias para buscas sistemáticas acadêmicas: Portal de Periódicos da CAPES, Portal de Teses e Dissertações da CAPES e Google Acadêmico. Para esses sistemas de busca, nossa avaliação revelou limitações no que diz respeito a interpretação correta de consultas, recuperação de dados e reprodutibilidade das pesquisas. A avaliação descrita neste relatório

\*Núcleo de Informática Aplicada à Educação (NIED) & Programa de Pós-Graduação Multiunidades em Ensino de Ciências e Matemática (PECIM), Universidade Estadual de Campinas (UNICAMP).

†Programa de Pós-Graduação Multiunidades em Ensino de Ciências e Matemática (PECIM), Universidade Estadual de Campinas (UNICAMP).

permite que pesquisadores possam i) analisar se um determinado sistema é adequado para seus requisitos de pesquisa sistemática e ii) utilizar critérios para realizar avaliações em outros sistemas de interesse.

**Palavras-chaves:** Revisão sistemática; Sistemas de busca acadêmica; Avaliação; Ensino.

### **Abstract**

Systematic reviews or literature reviews require the correct research strategy, using appropriate databases and efficient search systems. In the area of Education, there is a lack of publications on which search systems are suitable for systematic searches. This study shows an evaluation, with analysis of the search and retrieval qualities of 10 academic search systems widely used to carry out systematic searches by Brazilian researchers, considering the area of Education. The methodology for analyzing the search systems included 27 criteria divided into four categories: scope, search, search results and search reproducibility. This report shows the analysis of the CAPES Journal Portal, SciELO.br, Pubmed, Scopus, Web of Science (WoS), BDTD, LILACS, CAPES Theses and Dissertations Portal, Google Scholar and ERIC. The results show that while some systems can be recommended almost without limitation, others have failed important tests, which limit their suitability for systematic searches. The results indicate that three of the ten systems evaluated are not suitable as primary sources for systematic academic searches: CAPES Journal Portal, CAPES Theses and Dissertations Portal and Google Scholar. For these search systems, our evaluation revealed limitations regarding the correct interpretation of queries, data retrieval and reproducibility of searches. The evaluation described in this report allows researchers i) to analyze whether a given system is suitable for their systematic research requirements and ii) to use criteria to perform evaluations on other systems of interest.

**Keywords:** Systematic review; Academic search systems; Evaluation; Education.

# 1 Introdução

O crescimento constante das publicações de pesquisas científicas torna cada vez mais difícil e trabalhoso buscar e acompanhar as inovações de uma determinada área. O estudo e a síntese de evidências feito na forma de revisões sistemáticas é muito usual para resumir as evidências existentes sobre um assunto de pesquisa, identificar lacunas de pesquisa e assim poder guiar novas investigações (KITCHENHAM; CHARTERS, 2007).

A maioria das pesquisas começa com uma revisão da literatura. Esses estudos são baseados em pesquisas sistemáticas e avaliação crítica, visando produzir uma síntese imparcial das evidências sobre um determinado problema ou hipótese de pesquisa. Entretanto, se essa revisão não for completa e justa, será de pouco valor científico (KITCHENHAM; CHARTERS, 2007).

Uma revisão sistemática deve possuir uma estratégia de busca que permita replicação e avaliação por pares. O pesquisador ou a equipe de pesquisa que conduz a revisão deve ser capaz de realizar buscas online usando um conjunto adequado de sistemas, o que permitirá aos pesquisadores identificar todas as pesquisas disponíveis de maneira imparcial e analisar as mais relevantes. A busca nos bancos de dados é uma importante etapa da pesquisa acadêmica.

Atualmente, é muito difícil que um trabalho acadêmico seja feito sem o uso de sistemas de busca. Os resultados que os pesquisadores terão com suas consultas aos bancos de dados de pesquisa acadêmica dependem das funcionalidades dos sistemas de busca. Eles influenciam significativamente na recuperação de documentos e na precisão da busca. A capacidade dos sistemas de busca de recuperar resultados de maneira eficaz e eficiente determina sua adequação em pesquisas sistemáticas. Um sistema de busca adequado fornece boa cobertura na área de interesse e permite ao pesquisador especificar uma consulta com precisão. Assim, faz-se importante analisar essas plataformas de busca, onde se obtêm milhões de informações em alguns segundos (GUSENBAUER; HADDAWAY, 2020).

O objetivo desta pesquisa é avaliar e descrever as funcionalidades de sistemas de busca frequentemente utilizados em revisões sistemáticas, por pesquisadores brasileiros, na área de Ensino, focando nas qualidades e limitações que podem influenciar as revisões.

A capacidade e a precisão dos sistemas de busca podem ajudar os pesquisadores a decidir sobre onde investir melhor seu tempo, pois as buscas em revisões sistemáticas podem se tornar extensas e exaustivas.

A análise que realizamos se concentra nos 27 critérios de avaliação definidos por Gusenbauer e Haddaway (2020). Os autores utilizaram os 27 critérios em uma minuciosa avaliação de 28 sistemas de busca utilizados para revisões sistemáticas e meta-análises em todas as áreas, o que incluiu: *ACM Digital Library*, *AMiner*, *arXiv*, *Bielefeld Academic Search Engine (BASE)*, *CiteSeerX*, *ClinicalTrials.gov*, *Cochrane Library*, *Digital Bibliography & Library Project (DBLP)*, *Directory of Open Access Journals (DOAJ)*, *EbscoHost*, *Education Resources Information Cen-*

ter (ERIC), Google Acadêmico, IEEE *Xplore Digital Library*, JSTOR, Microsoft Academic, OVID, ProQuest, Pubmed, ScienceDirect, Scopus, Semantic Scholar, Springer Link, *Transport Research International Documentation* (TRID), *Virtual Health Library*, Web of Science (WoS), *Wiley Online Library*, WorldCat e WorldWideScience.

Diferente de Gusenbauer e Haddaway (2020), nossa análise se concentra em sistemas de busca acadêmica usados por pesquisadores brasileiros, populares na área de Ensino, um tema que carece de evidências e estudos comparativos sobre a eficácia dos sistemas de busca. Os dez sistemas de busca avaliados foram: Portal de Periódicos da CAPES, SciELO.br, Pubmed, Scopus, Web of Science (WoS), BDTD (Biblioteca Digital Brasileira de Teses e Dissertações), LILACS (Literatura Latino-Americana e do Caribe em Ciências da Saúde), Catálogo de Teses e Dissertações da CAPES, Google Acadêmico e ERIC (*Education Resources Information Center*).

Este estudo verifica se um sistema permite que o pesquisador especifique com precisão uma pesquisa para recuperar tantos resultados relevantes quanto possível, com que eficiência os resultados da pesquisa podem ser recuperados e se os resultados da pesquisa podem ser reproduzidos com os mesmos métodos.

A questão investigativa desta pesquisa é: **Quão adequados e utilizáveis são os sistemas de busca acadêmica, comumente usados por pesquisadores brasileiros, para realizar buscas sistemáticas para síntese de evidências na área de Ensino?**

Com esta pesquisa, não pretendemos indicar um banco de informações bom ou um ruim, mas sim as limitações de um ou outro, onde os menos notáveis podem ser usados como meios de buscas suplementares.

## 2 Metodologia para análise dos sistemas de busca

Conforme mencionado na introdução, neste relatório, usamos os 27 critérios de Gusenbauer e Haddaway (2020) para avaliar 10 sistemas de busca acadêmica populares entre os pesquisadores brasileiros da área de Ensino. Os critérios foram descritos neste relatório de uma maneira facilitada e podem ser utilizados por pesquisadores para avaliação de outros sistemas de busca acadêmica.

O único critério que adaptamos foi o número 11. No relatório de Gusenbauer e Haddaway (2020), esse critério diz respeito ao suporte de caracteres especiais, importantes para interpretar chinês e cirílico. Em nosso caso, testamos se o buscador retorna resultados especificamente em português, inglês e espanhol. O único critério excluído foi número 6, pois ele diz respeito a opções de pesquisa fornecidas pelo sistema de busca para um determinado banco de dados e, no caso da área de Ensino, é mais usual considerar apenas as coleções principais (*main collections*).

Os critérios e a metodologia de análise propostos por Gusenbauer e Haddaway (2020) foram criados com base em requisitos estipulados nas orientações de conduta emitidos por três instituições renomadas, dedicadas a garantir e elevar a qualidade da síntese de evidências na academia, como é o caso das revisões sistemáticas da literatura: Cochrane (2025), *The Campbell Collaboration* (CAMPBELL, 2025) e *Collaboration for Environmental Evidence* (CEE, 2025). As orientações PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) (PRISMA, 2020) e ROSES (*Reporting standards for Systematic Evidence Syntheses*) (ROSES, 2025) não foram incluídas no estudo, pois são mais concentradas em orientações sobre relatórios em vez de conduta.

As orientações para revisões sistemáticas referem-se a três requisitos de qualidade (GUSENBAUER; HADDAWAY, 2020):

1. deve ser possível identificar todos os registros relevantes, que atendem a critérios de elegibilidade;
2. a pesquisa deve ser transparente, ter um processo de pesquisa explícito e documentado, que permita atualizar ou replicar a pesquisa.
3. a pesquisa deve ser reproduzível, empregando os mesmos métodos com um ou mais sistemas de busca. Se a mesma consulta levar aos mesmos resultados de pesquisa, a pesquisa será considerada reproduzível.

Os 27 critérios foram divididos entre **necessários** (que precisam constar no sistema de busca) e **desejados** (que dependem dos anseios e objetivos de quem realiza a pesquisa). Os critérios necessários dizem respeito aos três requisitos de qualidade supramencionados. A descrição dos critérios é apresentada na próxima seção.

## 2.1 Critérios e procedimentos para análise

Os critérios utilizados para a análise dos sistemas de busca acadêmica foram divididos nas categorias: abrangência, busca, resultados da busca e reprodutibilidade da busca.

### 2.1.1 Critérios sobre Abrangência

Os critérios #1 ao #5 são referentes à abrangência. Um buscador menor com relação ao tamanho da sua base de dados, mas específico de uma área, pode trazer resultados de busca mais significativos em comparação a um buscador multidisciplinar. Isto é, mais nem sempre significa melhor. Todos os aspectos relacionados a abrangência foram caracterizados como aspectos desejados e não necessários.

**Critério #1 abrangência do conteúdo (desejável)** Refere-se ao tema acadêmico predominante no sistema de busca (Educação, Ensino, Ciências Sociais, Medicina, Multidisciplinar, etc).

**Critério #2 tamanho (desejável)** Refere-se ao número de registros na base de dados usada pelo sistema de busca.

**Critério #3 tipo de registro (desejável)** São os tipos de registros acadêmicos oferecidos (artigos de periódicos, relatos de experiência, resenhas, resumos de livros, anais de eventos, teses, dissertações, dentre outros).

**Critério #4 retrospectiva (desejável)** É a data do registro mais antigo presente no banco de dados.

**Critério #5 acesso aberto (desejável)** Refere-se ao uso de direitos autorais abertos. Os buscadores foram classificados de acordo com a predominância do tipo de direito autoral dos registros: se a grande maioria dos registros é de acesso pago/fechado, então o sistema de busca é classificado como **proprietário**; se a grande maioria é de acesso aberto, o buscador é classificado como **aberto**; se não há predominância de registros abertos ou fechados, então o buscador é **misto**.

### 2.1.2 Critérios para Busca

Os critérios #6 ao #23 são referentes à busca sistemática, utilizados para obter maior precisão ou maior número de resultados. Estes critérios abrangem detalhes no que se refere ao tipo de busca e como obter melhores resultados com base em alguns aspectos.

**Critério #6 vocabulário controlado (desejável)** Esse critério diz respeito a opções de pesquisa oferecidas para determinado banco de dados, no caso de bancos de dados que são acessados através de plataformas mais amplas, por exemplo, Web of Science e BVS (Biblioteca Virtual de Saúde, utilizada para acessar a base de dados LILACS). As opções disponíveis muitas vezes diferem entre bancos de dados. De acordo com Gusenbauer e Haddaway (2020) os bancos de dados nas ciências sociais tendem a não ser tão completamente indexados quanto os da medicina, por exemplo (p. 193). Nesta pesquisa, apenas as coleções principais (*core* ou *main collections*) usuais na área de Ensino, foram avaliadas. Por isso, esse critério foi desconsiderado.

**Critério #7 Campos para precisão da consulta (necessário  $\geq 5$ )** Refere-se a quais campos de busca estão disponíveis, visando obter um resultado mais preciso. Os códigos de campo permitem detalhar quais partes dos documentos devem ser pesquisadas. A disponibilidade de códigos de campo foi considerada um critério necessário, pois o pesquisador deve poder especificar onde as informações solicitadas estão localizadas nos registros dos documentos (título, resumo, palavras-chave, autor, dentre outros). Em geral, quanto mais

códigos de campo estiverem disponíveis melhores serão as chances fazer uma busca com alta precisão.

**Critério #8 Busca no texto completo (desejável)** Em alguns casos há a necessidade de buscar informações no texto na íntegra para identificar algo em específico.

**Critério #9 Tamanho máximo da equação de busca (necessário  $\geq 25$ )** Refere-se a quantos termos cabem na busca. Caso seja necessário usar uma equação de busca maior do que o suportado pelo sistema, é preciso dividir as equações. Esta prática pode, no entanto, ser trabalhosa e propensa a erros.

**Critério #10 tempo de resposta (desejável  $\leq 5s$ )** Refere-se ao tempo que o buscador leva para fazer uma busca que daria numerosos resultados.

**Critério #11 idioma dos registros (desejável)** Trata-se de verificar se os buscadores trazerem resultados em português, inglês e espanhol.

**Critério #12 ao #15 uso de booleanos (necessário)** Refere-se ao uso das palavras-chave booleanas AND, OR e NOT na equação de busca. É um dos testes mais importantes. Foram utilizados seis termos combinados de maneiras diferentes. Para buscas com aumento de OR espera-se que o número de registros encontrados aumente ou se mantenha constante; para buscas com aumento de AND espera-se uma redução na quantidade de registros encontrados ou que se mantenha constante; a mesma redução é esperada para NOT. É importante observar que a maneira de combinar os booleanos pode mudar de buscador para buscador.

**Critério #16 consultas literais (desejável)** Trata-se de verificar se os buscadores corrigem automaticamente palavras erradas (como esta).

**Critério #17 caractere curinga (desejável)** É desejável que o buscador ofereça a possibilidade de usar caracteres curinga. Os mais usuais são:

\* para truncamento à direita

? para caractere único

\$ para nenhum ou um caractere

**Critério #18 frase exata (necessário)** O buscador deve oferecer a possibilidade de fazer buscas com frases inteiras entre aspas. Os testes consistem em verificar se os resultados com e sem aspas diferem. É um critério necessário, pois é comum que os pesquisadores procurem por termos com combinações de palavras específicas.

**Critério #19 parênteses (necessário)** É importante que o buscador possa lidar com diferentes posicionamentos de parênteses para selecionar um grupo individual de termos que são relacionados.

**Critério #20 filtros para refinamento pós-busca (desejável)** Após uma busca ser realizada, há a possibilidade de refinar a pesquisa. Quanto mais poderosas as opções de filtro para o pós-consulta, maior a precisão de uma determinada consulta.

**Critério #21 citação avançada (desejável)** Trata-se do sistema de busca ser capaz de citar registros específicos de interesse. A lógica é que registros encontrados na busca, citam outros registros relevantes (mas que não foram retornados na busca). Através do relacionamento entre registros, o conjunto de resultados relevantes pode ser aumentado para além do que poderia ter sido encontrado apenas por meio da consulta.

**Critério #22 campos de busca avançada (desejável)** Os campos de entrada de busca avançada permitem aos pesquisadores compor as equações de pesquisa mais facilmente.

**Critério #23 ajuda (desejável)** É desejável que exista algum botão de ajuda ou outra opção na interface com instruções para ajudar o pesquisador na realização da busca.

### 2.1.3 Critérios sobre resultados da busca

São critérios que têm por finalidade facilitar o trabalho dos pesquisadores depois que a busca foi realizada.

**Critério #24 número máximo de acessos ao último registro (necessário)** É necessário que o sistema de busca ofereça uma maneira facilitada de acessar a última página de um grande número de resultados (maior ou igual a 1.000).

**Critério #25 download em massa (desejável)** É desejável que os resultados possam ser exportados em diferentes formatos, como BibTex, CSV, RIS, dentre outros. Trata-se de uma funcionalidade muito útil quando se deseja usar ferramentas para gerenciamento de referências, tais como Zotero, Mendeley, etc.; ou quando se deseja importar os resultados para planilhas. Os testes se concentraram em verificar se os registros podem ser baixados de uma só vez, envolvendo uma única seleção, e para quais formatos eles podem ser exportados.

### 2.1.4 Critérios sobre reprodutibilidade da busca

É preciso que a busca retorne os mesmos resultados quando executada em momentos e locais diferentes.

**Critério #26 reprodutibilidade da pesquisa em tempos diferentes (necessário)** É muito importante que os pesquisadores possam reproduzir a mesma pesquisa, com as mesmas equações e filtros, em momentos diferentes.

**Critério #27 reprodutibilidade da pesquisa em locais diferentes (necessário)** Mesmo em locais diferentes, o resultado da busca deve ser o mesmo.

No Quadro 1, resumimos os critérios e procedimentos utilizados para a análise de cada critério.

## 2.2 Seleção dos sistemas de busca para avaliação

Os 10 sistemas de busca analisados neste relatório foram escolhidos com base em 49 pesquisas do tipo revisões sistemáticas realizadas em um intervalo de 10 anos, em âmbito nacional, na área de Ensino.

A busca foi efetivada no Portal de Periódicos da CAPES, no dia 27/10/2022, com recorte temporal de 2012 até 2021 (uma década), usando a equação "revisão sistemática" AND ensino.

Encontramos 66 resultados, mas como havia três duplicados, o resultado final foi de 63 artigos para análise.

Definimos alguns critérios para exclusão de artigos: o estudo não estar relacionado com a área de Ensino; não ser uma pesquisa do tipo revisão sistemática; pesquisa realizada diretamente nos sistemas de busca de revistas específicas; não informar quais bancos de informações foram usados. Para além dos critérios acima, caso algum trabalho não pudesse ser acessado, também seria excluído da sondagem.

Deste modo, dos 63 artigos iniciais, 14 foram excluídos de acordo com os critérios acima, restando 49 para a catalogação.

A próxima etapa foi a catalogação dos sistemas de busca usados nas revisões sistemáticas dos artigos. No Quadro 2 tem-se os sistemas de busca utilizados e o número de vezes que foram empregados nas revisões sistemáticas. Foram considerados apenas os sistemas de busca que apareceram em quatro ou mais revisões sistemáticas.

## 3 Resultados

Os testes nos 10 sistemas de busca foram realizados em fevereiro e março de 2024, sendo a última consulta realizada em 04 de março de 2024. A avaliação foi feita sob a perspectiva do pesquisador, conforme os 27 critérios resumidos no Quadro 1. Nas próximas subseções, a avaliação dos critérios para abrangência, busca, resultados da busca e reprodutibilidade da busca são relatadas.

Quadro 1: Procedimentos utilizados para analisar os critérios.

#	Critério	Procedimento	Desejável (D) Necessário (N)
1	Abrangência	Buscar a descrição do tipo de disciplinas acadêmicas no site oficial	D
2	Tamanho	Buscar o número máximo de registros na base de dados no site oficial	D
3	Tipo	Buscar uma descrição dos tipos de registros oferecidos	D
4	Retrospectiva	Verificar a disponibilidade de registros mais antigos	D
5	Acesso	Verificar o tipo de direitos de uso atribuídos aos registros	D
6	Vocabulário	Não se aplica	–
7	Campos de consulta	Verificar a existência de códigos de campo para especificar em quais partes do documento a busca será realizada	Necessário $\geq 5$
8	Texto completo	Verificar a existência da opção de pesquisa de texto completo	D
9	Tamanho da equação	Verificar o número de palavras na equação de busca	Necessário $\geq 25$
10	Tempo de resposta	Verificar o tempo de resposta para uma equação com 25 termos que retorna muitos resultados	Desejável $\leq 5s$
11	Filtragem de idioma	Verificar se o sistema de busca possui uma maneira de filtrar registros em português, inglês e espanhol	D
12	Booleano OR	Testar se o booleano OR funciona com 6 termos. A cada termo acrescentado, o total de registros deve aumentar ou se manter constante	N
13	Booleano AND	Testar se o booleano AND funciona com 6 termos. A cada termo acrescentado, o total de registros deve diminuir ou ficar constante	N
14	Booleano NOT	Testar se o booleano NOT funciona com 6 termos. A cada termo acrescentado, o total de registros deve diminuir ou ficar constante	D
15	AND e OR	Testar diferentes combinações de AND e OR	N
16	Consultas literais	Corrigir palavras com escrita errada. Verificar se o sistema recupera poucos registros para essas palavras	D
17	Caracteres “curinga”	Verificar se caracteres curinga estão disponíveis: * para truncamento à direita ? para caractere único \$ para nenhum ou um caractere	D
18	Frase exata	Verificar se o uso de aspas (“ ”) altera os resultados da pesquisa	N
19	Parênteses	Testar diferentes posições de parênteses, com booleanos AND e OR	N
20	Filtros pós-consulta	Verificar se os filtros que permitem o refinamento de uma consulta estão visíveis na interface de pesquisa	Desejável $\geq 5$
21	Citação+	Verificar se o sistema de busca indica artigos relacionados	D
22	Busca+	Verificar opções de busca avançada na interface	D
23	Ajuda	Verificar se está disponível ajuda para auxiliar o pesquisador na condução da sua busca	D
24	Acesso aos resultados	Verificar o acesso à última página e às páginas intermediárias de um grande conjunto de resultados ( $\geq 1.000$ resultados)	D
25	Download em massa	Verificar quantos registros podem ser baixados de uma só vez, com uma única seleção, e para quais formatos exportar	D
26	Reprodutibilidade tempo	Consulta repetida após alguns dias	N
27	Reprodutibilidade local	Consulta repetida em locais diferentes	N

Fonte: Baseado em Gusenbauer e Haddaway (2020).

Quadro 2: Sistemas de busca que apareceram nos 49 artigos de revisão sistemática.

Portal de busca	n <sup>o</sup> vezes	Endereço
Portal de Periódicos da CAPES	17	<a href="https://www.periodicos.capes.gov.br">https://www.periodicos.capes.gov.br</a>
SciELO.br	15	<a href="https://www.scielo.br">https://www.scielo.br</a>
Pubmed	11	<a href="https://pubmed.ncbi.nlm.nih.gov">https://pubmed.ncbi.nlm.nih.gov</a>
Scopus	10	<a href="https://www.scopus.com">https://www.scopus.com</a>
Web of Science (WoS)	10	<a href="https://www.webofscience.com">https://www.webofscience.com</a>
BDTD (Biblioteca Digital Brasileira de Teses e Dissertações)	9	<a href="https://bdtd.ibict.br">https://bdtd.ibict.br</a>
LILACS (Literatura Latino-Americana e do Caribe em Ciências da Saúde)	8	<a href="https://lilacs.bvsalud.org">https://lilacs.bvsalud.org</a>
Catálogo de Teses e Dissertações da CAPES	7	<a href="https://catalogodeteses.capes.gov.br">https://catalogodeteses.capes.gov.br</a>
Google Acadêmico	6	<a href="https://scholar.google.com">https://scholar.google.com</a>
ERIC ( <i>Education Resources Information Center</i> )	4	<a href="https://eric.ed.gov">https://eric.ed.gov</a>

Fonte: Produzido pelos autores.

### 3.1 Abrangência

O Quadro 3 mostra os resultados dos testes de abrangência (critérios #1 ao #5).

Entre os 10 sistemas de busca, verificamos que 7 são multidisciplinares, 2 têm o foco em ciências médicas ou biomédicas (PubMed e LILACS) e 1 é especializado em Educação e Ensino (ERIC).

Poucas informações referentes aos critérios #2 a #5 foram obtidas em descrições contidas nos sites oficiais. A maioria foi obtida por meio da realização de buscas testes. Por exemplo, para descobrir o número de registros (critério #2), fizemos buscas com palavras que aparecem na grande maioria das pesquisas, como “pesquisa”, “ciência”, “the” e “research”. Os tamanhos dos bancos de dados variaram de cerca de 800 mil (BDTD) até cerca de 77,8 milhões (Scopus) de registros.

O tipo de registro (critério #3) foi verificado por intermédio dos campos de busca avançada, onde apareciam os tipos de registro como opção. Os tipos de registro disponíveis variaram substancialmente, como pode ser visto no Quadro 3. No Google Acadêmico não está claro quais são os tipos de registros indexados.

Com relação ao critério #4 novamente, usamos termos gerais, como “the” e anotamos o registro mais antigo. O registro mais antigo variou entre 1793 (PubMed) e 1987 (Portal de Periódicos da CAPES), sendo que o Google Acadêmico apresentou valores incertos. Para uma busca pela palavra “pesquisa” entre os anos de 1400 e 1800, o Google Acadêmico retornou diversos registros confundindo o número do ano da publicação com o número de endereços físicos de instituições e pessoas. Além disso, retornou um resultado onde o ano de 1796 aparece no título do artigo, mas a publicação é de 2022.

Quadro 3: Resultados dos testes dos critérios #1 ao #5.

<b>Portal</b>	<b>#1</b>	<b>#2</b>	<b>#3</b>	<b>#4</b>	<b>#5</b>
Periódicos da CAPES	Multidisciplinar	~1,5 milhão	Artigos, conjuntos de dados, resenhas, livros, periódicos, gravações de vídeo, imagem, dissertações, bases de dados, mapas	1888	Misto
SciELO.br	Multidisciplinar	~1,1 milhão	Artigo, artigo de revisão, relato de experiência, editorial, comunicação rápida, carta, resenha de livro, correção, resumo, comunicado de imprensa	1914	Acesso aberto
PubMed	Ciências médicas e biomédicas	~18 milhões	Artigos	1793	Misto
Scopus	Multidisciplinar	~77,8 milhões	Artigo de journal, artigo de conferência, editorial	1823	Proprietário
Web of Science	Multidisciplinar	~60 milhões	Artigos, cartas, resumos, crítica literária, capítulo de livro, discussão, retratação de publicação, bibliografia, reimpressão, análise de hardware, etc.	1898	Proprietário
BDTD	Multidisciplinar	~800 mil	Teses e Dissertações	1862	Acesso aberto
LILACS	Saúde da América Latina e Caribe	~1 milhão	Estudos clínicos, sínteses, revisões, guias, relatório técnico, publicações governamentais, teses e dissertações	1906	Acesso aberto
Banco de Teses da CAPES	Multidisciplinar	~1,5 milhão	Teses e Dissertações	1987	Acesso aberto
Google Acadêmico	Multidisciplinar	13 milhões	Incerto	Incerto	Misto
ERIC	Educação e Ensino	~1,8 milhão	Artigos, relatos de experiência, relatórios, artigos de opinião, discursos, guias, questionários, dissertações e teses	1839	Proprietário

Fonte: Produzido pelos autores.

Nossa amostra examinou 4 bancos de dados com recursos abertos (*Open Access*), 3 com recursos mistos e 3 sistemas de busca majoritariamente proprietários. Scopus, Web of Science e ERIC foram classificados como proprietários, mas permitem selecionar apenas os registros com acesso aberto na busca. O Portal de Periódicos da CAPES foi classificado como misto, mas também permite selecionar apenas registros com acesso aberto. No Pubmed, a informação sobre acesso aberto fica junto do registro retornado na busca. No Google Acadêmico, essa informação não está disponível.

## 3.2 Buscas

Os Quadros 4 e 5 mostram os resultados dos testes referentes à busca, sendo os critérios #7 ao #15 apresentados no Quadro 4 e os critérios #16 ao #23 no Quadro 5.

Quadro 4: Resultados dos testes dos critérios #7 ao #15.

Portal	#7	#8	#9	#10	#11	#12	#13	#14	#15
Periódicos CAPES	Não (=3)	Não	Sim	7,3s	Sim	Sim	Sim	Sim	Sim
SciELO.br	Sim	Não	Sim	1s	Não	Sim	Sim	Não	Sim
Pubmed	Sim	Sim	Sim	1,4s	Sim	Sim	Sim	Sim	Sim
Scopus	Sim	Não	Sim	2,8s	Sim	Sim	Sim	Sim	Sim
WoS	Sim	Não	Sim	1,5s	Sim	Sim	Sim	Sim	Sim
BDTD	Sim	Não	Sim	2s	Sim	Sim	Sim	Sim	Sim
LILACS	Sim	Não	Sim	8s	Sim	Sim	Sim	Sim	Sim
Teses CAPES	Não	Não	Sim	ERRO 8,3s	Não	Sim	Sim	Sim	Sim
Google Acadêmico	Não	Sim	Sim	0,11s	Não	Sim	Sim	Não	Não
ERIC	Sim	Não	Sim	0,9s	Não	Sim	Sim	Sim	Sim

Fonte: Produzido pelos autores.

O critério #7 diz respeito aos campos que tornam possível especificar em quais partes do documento a busca será realizada (título, abstract, palavras-chave, etc). O Google Acadêmico e o Catálogo de Teses da CAPES são os únicos que não permitem ao pesquisador especificar quais partes do registro serão consideradas na busca. No Google Acadêmico, nas configurações avançadas, é possível especificar se a busca pode ser realizada no texto completo ou no título. Já no Catálogo de Teses da CAPES, não está claro em quais partes do documento a busca é realizada. O Portal de Periódicos da CAPES oferece apenas três opções que permitem especificar em qual local do registro a busca será realizada (título, autor e assunto). Os demais sistemas de busca, possuem cinco campos ou mais.

Sobre o critério #8 (busca no texto completo), os únicos que permitem são o Google Acadêmico e o PubMed. Este último possui a opção “text word” na busca avançada.

Para testar o critério #9, utilizamos uma equação com 25 termos usuais na área de Ensino,

ligadas pelo conector OR. Todos os sistemas de busca processaram a consulta com os 25 termos.

pesquisa OR ciências OR educação OR ensino OR artigo OR resultado OR  
objetivo OR ambiente OR matemática OR física OR saúde OR química OR  
biologia OR natureza OR TIC OR conceito OR tempo OR análise OR relação  
OR campo OR área OR desafio OR metodologia OR conclusão OR documento

Para testar o critério #10 (tempo de resposta), utilizamos a mesma equação mostrada acima, com os 25 termos conectados pelo OR. Os tempos variaram substancialmente – o mais rápido foi o Google Acadêmico e o mais lento foi o Catálogo de Teses da CAPES<sup>1</sup>. Este último retornou erro na primeira tentativa; apenas na segunda tentativa foi possível medir o tempo aproximado de resposta.

Sobre o critério #11 (múltiplas línguas), em seis sistemas de busca é possível selecionar entre diversos idiomas, inclusive português, inglês e espanhol. No PubMed, por exemplo, é possível fazer a busca por artigos escritos em diversas línguas adicionando um filtro específico. O Portal de Periódicos da CAPES tem opções para retornar resultados em inglês, francês e alemão; para resultados em português e espanhol, a opção “qualquer idioma” fica acionada como padrão e há uma opção de filtragem após a busca. No SciELO.br, as opções são Português, Francês e Russo – para Inglês e Espanhol é preciso selecionar “Todos”. O Catálogo de Teses da CAPES não possui opção de idioma, aparentemente, mostra registros apenas em português. O BDTD, assim como o Catálogo de Teses da CAPES, possui dissertações e teses defendidas no Brasil, mas apresenta um filtro pós-consulta para possibilitar filtrar o idioma no qual a tese ou dissertação está escrita. No Google acadêmico, há apenas duas opções referentes a idioma (qualquer idioma ou o idioma para o qual o browser está configurado – em nosso caso, as opções eram “qualquer idioma” e “português”). Os artigos no ERIC são em inglês; há uma opção sobre idioma na busca avançada, mas essa opção não funcionou em nossos testes.

Para testar o critério #12 (booleano OR), começamos com a palavra “pesquisa” e acrescentamos mais 5 palavras gradativamente, ligadas pelo conector OR. O mesmo foi feito para testar o AND e o NOT (critérios #13 e #14).

pesquisa OR ciências OR educação OR ensino OR artigo OR resultado  
pesquisa AND ciências AND educação AND ensino AND artigo AND resultado  
pesquisa NOT ciências NOT educação NOT ensino NOT artigo NOT resultado

No ERIC, como os artigos estão em inglês, mudamos as palavras para

research OR science OR education OR teaching OR article OR result  
research AND science AND education AND teaching AND article AND result  
research NOT science NOT education NOT teaching NOT article NOT result

---

<sup>1</sup>As medidas são aproximadas, foram feitas com a utilização de um cronômetro.

A maioria dos testes retornaram resultados conforme o esperado. Isto é, ao acrescentar termos ligados pelo conector OR os resultados aumentavam, ao acrescentar termos ligados pelos conectores AND e NOT os resultados diminuam. As únicas exceções foram o Google Acadêmico e o SciELO.br com o conector NOT. Ao realizar uma busca com palavras em português utilizando o conector NOT, o Google Acadêmico fez confusão com NOTA e NA. Para uma busca com palavras em inglês ele, aparentemente, reconheceu o conector NOT. No entanto, ao mudar a quantidade de palavras da busca, verificou-se que os resultados ora aumentavam, ora diminuam, o que nos leva a concluir que o Google Acadêmico não reconhece o NOT. No SciELO.br, os resultados com AND e OR foram satisfatórios, mas o comportamento com o NOT é o mesmo do conector AND, o que nos leva a concluir que o SciELO.br não reconhece o NOT.

É importante observar que a maneira de escrever os booleanos pode mudar de buscador para buscador. Por exemplo, o Scopus troca o NOT por “AND not AND”. O Google Acadêmico, aparentemente, aceitou o OR em português. A mesma equação foi processada usando OU (pesquisa OU ciências OU educação OU ensino OU artigo OU resultado); mas, estranhamente, os resultados retornados foram diferentes da equação executada com OR. Sugerimos que os pesquisadores realizem testes antes de efetivarem suas buscas.

Para o critério #15 (combinação de AND e OR), usamos a seguinte equação:

educação OR ensino AND distância

Todos os sistemas funcionaram, exceto o Google Acadêmico. Para esse exemplo particular, ele confunde o AND com A e “deduz” que a busca é educação OR ensino A distância. O Google Acadêmico confunde os booleanos com termos ou palavras usuais, no exemplo ele confundiu ensino AND distância com ensino A distância.

É importante observar que a ordem de execução pode ser diferente a depender do sistema de busca. Para o exemplo acima, o Portal de periódicos da CAPES executa primeiro o AND, já o Scopus executa o que encontra primeiro (ele executa o OR primeiro).

Para o critério #16 (se identifica palavras escritas errado), o Pubmed mostrou uma mensagem dizendo que não houve resultados para a busca com a palavra errada e retornou resultados para a busca com outra palavra que ele julgou ser a correta. O Google Acadêmico teve um comportamento similar. O Scopus e o BDTD, sugerem uma busca com a palavra correta, mas não executam a busca. Os demais são sugerem que a palavra está errada, apenas retornam a busca com nenhum resultado. É importante observar que o corretor ortográfico do navegador pode ajudar (o corretor ortográfico do Google Chrome avisava quando a palavra estava escrita de forma errada).

Para o critério #17 (caractere curinga), usamos a palavra menina com as variações `menin?`, `menina$` e `menin*`.

Quadro 5: Resultados dos testes dos critérios #16 ao #23.

Portal	#16	#17	#18	#19	#20	#21	#22	#23
Periódicos CAPES	Não	Não	Sim	Sim	Sim	Não	Sim	Não
SciELO.br	Não	Sim	Sim	Sim	Sim	Não	Sim	Não
Pubmed	Sim	Não	Sim	Sim	Sim	Sim	Sim	Sim
Scopus	Sim	Não	Sim	Sim	Sim	Sim	Sim	Sim
WoS	Não	Sim	Sim	Sim	Sim	Sim	Sim	Não
BDTD	Sim	Não	Sim	Sim	Sim	Sim	Sim	Não
LILACS	Não	Sim	Sim	Sim	Sim	Não	Sim	Não
Teses CAPES	Não	Não	Sim	Não	Sim	Não	Não	Não
Google Acadêmico	Sim	Não	Sim	Não	Não <sup>2</sup>	Não	Sim	Não
ERIC	Não	Não	Sim	Sim	Sim	Não	Não	Sim

Fonte: Produzido pelos autores.

**menin?**: espera-se que os resultados aumentem com relação a menina, pois abrange menina, menino, menino.

**menina\$**: espera-se que aumente com relação a menina, pois abrange menina e meninas.

**menin\***: espera-se que os resultados aumentem com relação aos anteriores, pois engloba menina, meninas, menino, menino, menina, menino, meningite, etc.

Google Acadêmico e ERIC não tiveram o comportamento esperado para nenhum dos caracteres curinga, nem mesmo para o caractere de truncamento \* (que é o mais comum). No Google Acadêmico, em todos os testes, o resultado da busca diminuiu ao invés de aumentar. No ERIC, os testes foram realizados com a palavra **create**. Os resultados não se alteraram ao acrescentar os curingas (para o teste **creat\***, o resultado diminuiu).

O Portal de Periódicos da CAPES e o Catálogo de Teses da CAPES reconhecem apenas o caractere curinga para truncamento \*. Nos testes com os outros dois, o comportamento não foi como esperado.

O Pubmed, o Scopus e o BDTD, parecem não reconhecer o caractere \$. O caractere \* pode ser utilizado para substituí-lo.

O WoS, o SciELO.br e o LILACS reconheceram os 3 caracteres curinga testados.

Os testes mostraram uma grande variação dos sistemas de busca com relação ao uso de caracteres curinga. Sugere-se que os pesquisadores testem se esses caracteres funcionam adequadamente antes de usá-los.

Para o critério #18 (frase exata) usamos as equações com e sem aspas:

"Educação a distância"

"Educação a distância" OR "Ensino a distância": espera-se aumento do resultado.

"Educação a distância" AND "Ensino a distância": espera-se diminuição do resultado.

Todos os sistemas de busca se comportaram conforme o esperado.

O posicionamento dos parênteses (critério #19) é muito importante nas buscas sistemáticas, pois eles influenciam na lógica da busca e em quais partes serão executadas primeiro. Para o critério #19, fizemos os seguintes testes:

(educação OR ensino) AND distância  
educação OR (ensino AND distância)

Como a mudança de posição dos parênteses influencia na lógica da equação, espera-se que o resultado da execução das duas equações seja diferente. Para esse exemplo, os únicos a ignorar a presença dos parênteses foram o Catálogo de Teses da CAPES e o Google Acadêmico. Em ambos, com ou sem os parênteses, os resultados foram os mesmos, sendo que o Google Acadêmico continuou a confundir ensino AND distância com ensino A distância, independente da ordem dos parênteses.

Ainda no critério #19, testamos equações com diversos elementos juntos (AND, OR, parênteses e aspas):

(Conceituação OR Conceito OR Concepção OR Fundamento) AND  
("Educação a Distância" OR "Ensino a Distância" OR EaD)

No ERIC:

(Concept OR Conception) AND  
("Distance Education" OR E-learning)

Não houve confusão por parte do Google Acadêmico dessa vez. O único que não se comportou conforme o esperado foi o Catálogo de Teses da CAPES. Esse sistema de busca não relacionou o resultado do que estava dentro dos parênteses com o AND. Foi necessário quebrar a equação; fazendo todas as combinações possíveis das palavras dentro do primeiro par de parênteses com as palavras do segundo par de parênteses. Dividir a equação é uma prática mais trabalhosa e propensa a erros. Para execução no sistema de busca do Catálogo de Teses da CAPES, a equação acima ficou deste jeito:

Conceituação AND "Educação a Distância"  
Conceituação AND "Ensino a Distância"  
Conceituação AND EaD  
Conceito AND "Educação a Distância"  
Conceito AND "Ensino a Distância"  
Conceito AND EaD  
Concepção AND "Educação a Distância"  
Concepção AND "Ensino a Distância"  
Concepção AND EaD  
Fundamento AND "Educação a Distância"  
Fundamento AND "Ensino a Distância"  
Fundamento AND EaD

Em todos os sistemas, os resultados da pesquisa podem ser refinados por meio de filtros pós-consulta (critério #20). Alguns sistemas, como PubMed, Scopus e WoS oferecem mais de 30 opções diferentes para filtrar seus bancos de dados altamente estruturados. O ERIC oferece filtros com descritores específicos das áreas de Educação e Ensino. Uma particularidade do ERIC que chamou a atenção foi a seleção do ano das publicações. Não é possível especificar o ano da busca, com precisão, utilizando apenas a interface de busca. As opções apresentadas na interface foram: *In 2024*, *since 2023*, *since 2020 (last 5 years)*, *since 2015 (last 10 years)*, *since 2005 (last 20 years)*. É necessário usar as palavras-chaves *pubyearmin* e *pubyearmax* para especificar um período de tempo (em anos) que não esteja visível na interface.

O Google Acadêmico é o mais limitado com relação aos filtros pós-consulta, oferece apenas 4 opções. Notadamente, um filtro muito importante para a realização de revisões sistemáticas – a seleção do tipo de registro (artigo de periódico, artigo de revisão, tese/dissertação, anais de eventos, dentre outros) – está presente no Google Acadêmico de uma maneira muito simplificada. Ele mostra apenas as opções de “artigo de revisão” ou “todos os tipos”, inviabilizando a filtragem de artigos de periódicos, anais e outros tipos de registro importantes em trabalhos acadêmicos.

A pesquisa de citação avançada (critério #21) estava disponível em 4 sistemas de busca (PubMed, Scopus, WoS e BDTD). No SciELO.br existe uma opção de citação avançada, mas está vinculada ao Google Acadêmico.

Os únicos sistemas de busca que não oferecem campos para busca avançada na interface (critério #22) são Catálogo de Teses da CAPES e ERIC.

Sobre a opção de ajuda (critério #23), seis portais ofereceram alguma forma de ajuda, mas apenas 3 delas são boas para auxiliar os pesquisadores na condução sua busca (PubMed, Scopus e ERIC). O PubMed oferece *help* e *user guide*. No Scopus, a ajuda fica visível apenas antes de realizar a busca. O ERIC oferece um FAQ com tópicos relevantes para a busca acadêmica, na opção “*searching ERIC*”. O Portal de Periódicos da CAPES, oferece um menu com ajuda, mas

os tópicos não são pra ajudar o pesquisador na condução da sua busca, são sobre a entrada no portal, modalidades de acesso, dentre outros. O SciELO.br não oferece ajuda em nenhum momento. No LILACS não há ajuda, apenas um link no rodapé, onde o pesquisador pode preencher um formulário solicitando ajuda ou enviando sugestões. No Google Acadêmico, a ajuda também fica no rodapé e dá informações sobre a busca no Google de maneira geral, não ajuda no caso de especificidades de buscas acadêmicas. No Catálogo de Teses da CAPES, a ajuda é genérica e confusa. Por exemplo, na ajuda diz: “Para pesquisar dois termos, use +, por exemplo: Educação + Científica”.

Realizamos testes com **Educação + Científica**, **Educação AND Científica**, **Educação OR Científica**. Esperávamos que os resultados dos dois primeiros fossem iguais, mas os três resultados foram completamente diferentes. Portanto, nossos testes mostraram que a informação na Ajuda não está correta.

### 3.3 Resultados da busca

Os testes dos critérios #24 e #25 referem-se ao acesso e manipulação dos registros após a realização da busca. Os resultados estão no Quadro 6.

Quadro 6: Resultados dos testes dos critérios #24 e #25.

Portal	#24a	#24b	#25	Formatos exportados
Periódicos CAPES	Sim	Não	Sim	BibTex, RIS, Refworks, EndNote, EasyBib
SciELO.br	Sim	Sim	Sim	RIS, BibTex, CSV
Pubmed	Sim	Sim	Sim	PMID, texto (sumário e abstract), CSV e Pubmed-style
Scopus	Sim	Sim	Sim	CSV, RIS, BibTex
WoS	Sim	Sim	Sim	texto, RefWorks, RIS, BibTex, XLS (Excel), HTML, etc.
BDTD	Não	Não	Sim	CSV
LILACS	Sim	Não	Sim	RIS, CSV, texto (citação)
Teses CAPES	Não	Não	Não	–
Google Acadêmico	Sim	Não	Não	BibTex, EndNote, RefMan, RefWorks
ERIC	Não	Não	Sim	Medline/Pubmed (.nbib)

Fonte: Produzido pelos autores.

Com relação ao critério #24, foi testado se é permitido selecionar resultados por página (#24a) e se é possível ir até o último registro sem precisar passar pelos demais (#24b).

O número máximo de registros recuperáveis por página (critério #24a) varia muito entre os sistemas de busca. Quatro deles oferecem opções para selecionar a quantidade de registros por página (10, 30, 50, por exemplo). Pubmed, Scopus e LILACS permitem exibir mais que 50 itens por página. BDTD, Catálogo de Teses da CAPES e ERIC exibem um número fixo de registros por página (20 registros os dois primeiros e 15 registros o último). No Quadro 6, apenas os

sistemas que permitem exibir cinquenta ou mais registros por página foram classificados como satisfatórios e receberam o “Sim”.

No Portal de Periódicos da CAPES, não é possível chegar até a última página sem passar por todas as anteriores (critério #24b). No BDTD, LILACS, Catálogo de Teses da CAPES e ERIC é possível ir diretamente até a última página, mas não é possível ir diretamente a uma página específica. No Google Acadêmico, é possível acessar páginas intermediárias, mas não é possível ir até a última página sem percorrer algumas anteriores. No Quadro 6, apenas os sistemas que permitem ir diretamente a uma página específica, inclusive à última, foram classificados como satisfatórios e receberam o “Sim” (SciELO.br, PubMed, Scopus e WoS).

As opções de download em massa (critério #25) diferiram significativamente entre os sistemas de busca. Seis permitem o download de todos os resultados de uma só vez (SciELO.br, Pubmed, Scopus, WoS, BDTD, LILACS). Apenas um dos sistemas analisados faz o download apenas dos itens selecionados em uma mesma página (Portal de Periódicos da CAPES). ERIC e SciELO.br fazem a exportação em blocos de tamanho definido (no máximo 200 registros de cada vez). Dois dos sistemas examinados não forneciam suporte para exportação em massa de registros (Catálogo de Teses da CAPES e Google Acadêmico). Nas configurações do Google Acadêmico existe uma opção para exportação de registros, mas é feito individualmente para cada registro, não em massa. Os sistemas que não oferecem opção para exportação em massa não foram considerados satisfatórios nesse critério (Catálogo de Teses da CAPES e Google Acadêmico).

Conforme apresentado no Quadro 6, os formatos mais utilizados para exportação são CSV, BibTex e RIS. A exportação em massa e os formatos da exportação são importantes para a interoperabilidade dos resultados da busca com outros sistemas, que auxiliam na organização dos registros para análise e facilitam muito os trabalhos de revisão sistemática. O formato CSV é comumente utilizado em softwares de planilhas eletrônicas como Excel e LibreOffice Calc. O formato BibTex é utilizado no Latex, usual para a escrita de artigos, teses e dissertações. O formato RIS é muito comum em softwares gerenciadores de referências como Mendeley e Zotero.

### 3.4 Reprodutibilidade da busca

Para testar a reprodutibilidade da busca (critérios #26 e #27), foi utilizada a equação:

```
(Conceito OR Fundamento) AND  
("Educação a Distância" OR EaD)
```

Nas bases de dados com resultados majoritariamente em inglês (Pubmed, Scopus e WoS), foi utilizada a equação:

(Concept OR Conception) AND  
("Distance Education" OR e-learning)

No ERIC, conforme mencionado na seção anterior, nem sempre é possível especificar o ano da busca utilizando apenas a interface de busca. Foi necessário usar as palavras-chaves `pubyearmin` e `pubyearmax`. A equação de busca no ERIC foi a seguinte:

(Concept OR Conception) AND  
("Distance Education" OR E-learning)  
`pubyearmin:2022 pubyearmax:2023`

No Catálogo de Teses da CAPES, conforme percebemos na avaliação do critério #19 (seção 3.2), as equações contendo combinações de parênteses com conectores booleanos não funcionam apropriadamente. Por isso, foi necessário usar quatro equações:

Conceito AND "Educação a distância"  
Conceito AND EaD  
Fundamento AND "Educação a distância"  
Fundamento AND EaD

No Google Acadêmico foram testadas equações com conectores booleanos em português e inglês, pois ao utilizar os booleanos em inglês, ele sugeriu o uso em português:

(Conceito OU Fundamento) E ("Educação a distância" OU EaD)  
(Conceito OR Fundamento) AND ("Educação a distância" OR EaD)

Os resultados das buscas com os booleanos em português foram muito diferentes dos resultados com os booleanos em inglês. Anotamos no Quadro 7 os resultados da busca com os booleanos em português.

Foram considerados na busca dissertações, teses e artigos de periódico (inclusive artigos de revisão) publicados nos anos de 2022 e 2023. Três testes foram realizados, sumarizados no Quadro 7:

1. Teste 1: A primeira busca foi realizada em 22 de fevereiro de 2024, na Universidade.
2. Teste 2: Uma semana depois do teste 1, no dia 29 de fevereiro de 2024, a mesma busca, com o mesmo recorte temporal foi refeita, mas em uma localidade diferente, sem a utilização de VPN (sem um IP da Universidade).

3. Teste 3: Quatro dias depois do teste 2, a terceira e última busca foi realizada, no dia 04 de março de 2024, com dois pesquisadores ao mesmo tempo, em localidades diferentes, um dos pesquisadores em casa usando VPN e outro na Universidade.

No Google Acadêmico, ao avaliar o critério #20 (seção 3.2) já havíamos percebido que não é possível selecionar apenas teses, dissertações e artigos de revista. Para verificar a reprodutibilidade da busca no Google Acadêmico, selecionamos primeiramente “artigos de revisão” e, em seguida “qualquer tipo”.

Nos três testes realizados, os resultados foram os mesmos em 6 sistemas de busca (Periódicos CAPES, SciELO.br, PubMed, LILACS, Catálogo de Teses CAPES e ERIC). Esse é o resultado esperado, isto é, independente do momento e do local, as buscas devem retornar os mesmos resultados.

Certas plataformas demandam ter uma assinatura, em nosso caso Scopus e Web of Science, onde é necessário ter login e senha para realizar consultas. Essas plataformas podem apresentar resultados diferentes, dependendo na instituição através da qual o banco de dados é acessado. Isto é, o número de registros depende da assinatura feita pela instituição. Por exemplo, a coleção principal da Web of Science varia em escopo dependendo da assinatura. Outra questão é que uma instituição pode ter feito uma assinatura com cobertura desde 1996, e outra desde 2010. Nesses casos, a recuperação dos registros será diferente (GUSENBAUER; HADDAWAY, 2020, p.208). Essas variações mostram que os pesquisadores devem estar familiarizados com a assinatura da sua instituição e documentar isso em suas revisões para que elas sejam de fato reprodutíveis.

Quadro 7: Resultados dos testes dos critérios #26 e #27.

Portal	Teste 1	Teste 2	Teste 3	#26	#27
Periódicos CAPES	80 resultados (artigos)	80	80	Sim	Sim
SciELO.br	3 (article) 1 (review article)	3 1	3 1	Sim	Sim
Pubmed	170 resultados	170	170	Sim	Sim
Scopus	1.116 (article) 57 (review)	Não entra sem login	1.117 57	Sim	Sim
WoS	365 (artigos) 28 (review article)	Não entra sem login	366 28	Sim	Sim
BDTD	75 resultados (teses/dissertações)	84	84	Sim	Sim
LILACS	14 resultados	14	14	Sim	Sim
Teses CAPES	5 resultados Nenhum resultado 8 resultados 3 resultados Total = 16 resultados (teses/dissertações)	5 0 8 3 Total = 16	5 0 8 3 Total = 16	Sim	Sim
Google Acadêmico	84 (artigos de revisão) 5.000 (qualquer tipo)	82 4.630	86 4.730	Não	Sim
ERIC	605 (Journal Articles) 79 (Dissertations/Thesis)	605 79	605 79	Sim	Sim

Em nossos testes, o BDTD teve um aumento no número de registros no período de 1 semana (de 75 para 84). O Scopus e o Web of Science tiveram um pequeno aumento no período de 10 dias (aumentaram 1 registro cada). Trata-se de um crescimento natural, pois o conjunto de dados indexado em um banco de dados aumenta com a identificação e curadoria de novos registros. Os resultados tendem a aumentar, principalmente se o recorte temporal for próximo ao mês/ano vigente.

Nos testes do BDTD, Scopus e Web of Science, consultas repetidas produziram um conjunto maior de resultados do que a consulta inicial (critério #26). Isso deve ser levado em consideração na avaliação da reprodutibilidade do sistema de pesquisa, e não deve ser visto como uma falta de reprodutibilidade. Os pesquisadores devem, portanto, relatar a data em que as pesquisas foram realizadas para tornar compreensíveis as alterações na base de dados do sistema de pesquisa (GUSENBAUER; HADDAWAY, 2020).

Apenas o Google Acadêmico falhou na reprodutibilidade com relação ao critério #26 (realização da busca em momentos diferentes). Os resultados das buscas ora aumentavam e ora diminuam, como pode ser observado no Quadro 7. O conjunto de resultados diferiu em um caminho não explicável pelo crescimento natural do banco de dados.

Por fim, todos os sistemas de busca se comportaram conforme o esperado em nossa análise dos resultados recuperados em locais diferentes (critério #27). Os testes foram realizados pelos dois pesquisadores autores deste relatório, utilizando a mesma assinatura institucional, ao mesmo tempo, com ambos localizados no estado de São Paulo, em cidades diferentes (cerca de 50 km de distância).

## 4 Discussão

Os sistemas de busca analisados são diversos em suas funcionalidades e recursos. Os sistemas têm diferentes bancos de dados, métodos de indexação e apresentação de dados. Mecanismos de pesquisa na Web baseados em rastreadores (por exemplo, Google Acadêmico) funcionam de maneira diferente de bancos de dados bibliográficos que possuem um catálogo com registros específicos de resultados de pesquisa.

Na avaliação realizada, um sistema de busca pode ser adequado como recurso principal ou suplementar. Para ser considerado um sistema de busca principal, é preciso atender a todos os critérios considerados necessários no Quadro 1. Caso contrário, o sistema de busca foi considerado suplementar. Recursos suplementares podem ser usados adicionalmente a um recurso principal por suas qualidades específicas que poderiam recuperar registros adicionais e melhorar as evidências. Sendo assim, os pesquisadores precisam avaliar individualmente o quão importantes os critérios são para sua pesquisa sistemática. Outros trabalhos na literatura que reportam avaliação da qualidade de sistemas de busca também utilizaram a distinção entre re-

cursos principais e secundários para revisões sistemáticas (BRAMER et al., 2013; HADDAWAY et al., 2015; GUSENBAUER; HADDAWAY, 2020).

Critérios de qualidade específicos de pesquisadores não foram considerados na classificação de recursos principais e suplementares. Os critérios foram incluídos nos testes para informar os pesquisadores sobre funcionalidades dos sistemas de busca que são úteis ou importantes, mas sem considerar particularidades para cumprir os requisitos de qualidade de revisões sistemáticas de grupos de pesquisa específicos. Quanto mais os pesquisadores estiverem cientes das funcionalidades de um sistema de busca, melhor poderão otimizar sua estratégia de busca. Assim, os testes realizados neste estudo avaliaram funcionalidades para que os pesquisadores tenham uma visão da funcionalidade dos sistemas de busca para cada critério avaliado. Dessa forma, os pesquisadores podem consultar rapidamente critérios de qualidade individuais e refletir se um sistema de busca de interesse oferece funcionalidades adequadas às suas necessidades.

De acordo com os testes realizados, sete dos dez sistemas de busca analisados são adequados para a síntese de evidências na forma de revisões sistemáticas na medida em que atenderam a todos os requisitos considerados necessários (SciELO.br, PubMed, Scopus, Web of Science, BDTD, LILACS e ERIC).

Os sistemas de busca que não atenderam aos critérios, não são considerados adequados para uso como fontes principais de pesquisa para revisões sistemáticas (Portal de Periódicos da CAPES, Catálogo de Teses e Dissertações da CAPES e Google Acadêmico). Para esses sistemas de busca, os testes revelaram limitações no que diz respeito a interpretação correta de consultas (equações) pelo sistema, recuperação de dados, e a reprodutibilidade das pesquisas.

Em nossa avaliação, o Google Acadêmico foi considerado inadequado para revisão primária por ter sido reprovado em quatro critérios considerados necessários (critérios #7, #15, #19 e #26): Esse sistema de busca confunde os resultados dependendo das palavras conectadas pelos booleanos AND e OR (critério #15) - para o exemplo feito ao testar o critério #15, ele confundiu **ensino AND distância** com **ensino A distância**. A mesma confusão aconteceu mesmo com o uso dos parênteses (critério #19). Além disso, não é possível especificar quais partes do registro serão consideradas na busca (título, abstract, palavras-chave, etc) (critério #7) e não está claro quais são os tipos de registro indexados (critério #3): teses, dissertações, artigos de revista, artigos de conferência, editorial, etc. O Google Acadêmico deixou de entregar resultados replicáveis em momentos diferentes das buscas (critério #26). O número de resultados diferiu em um caminho não explicável pelo crescimento natural do banco de dados. Outros trabalhos na literatura também consideram o Google Acadêmico como uma fonte complementar para a síntese de evidências (HADDAWAY et al., 2015; GUSENBAUER; HADDAWAY, 2020).

Em nossa avaliação, o Catálogo de Teses e Dissertações da CAPES também foi considerado inadequado para revisão primária, pois foi reprovado em dois critérios considerados necessários (critérios #7 e #19): Não é possível especificar quais partes do registro serão consideradas na busca (título, abstract, palavras-chave, etc) (critério #7). O posicionamento dos parênteses

é muito importante nas buscas sistemáticas, pois podem influenciar na lógica da busca. O Catálogo de Teses e Dissertações da CAPES ignora a presença dos parênteses (critério #19). Quando o uso dos parênteses é necessário, é preciso quebrar a equação para que o sistema possa processá-la, conforme mostrado em exemplos nas seções 3.2 e 3.4.

O Portal de Periódicos da CAPES também foi considerado inadequado para revisão primária em nossa avaliação, pois ele reprovou no critério #7. Esse sistema de busca oferece apenas três opções que permitem especificar em qual local do registro a busca será realizada (título, autor e assunto). Não é possível realizar uma busca no abstract, o que limita e, em alguns casos, inviabiliza, o uso desse sistema para a síntese de evidências.

É surpreendente que três dos sistemas de pesquisa que examinamos (SciELO.br, Catálogo de Teses da CAPES e Google Acadêmico) tenham algum problema com consultas booleanas e/ou com o uso de parênteses (critérios #14, #15 e #19). Isto é lamentável porque a pesquisa booleana é eficaz, especialmente para estratégias de pesquisa sistemática (HJØRLAND, 2015).

Outra constatação alarmante percebida em nossa avaliação são as limitações dos sistemas de busca mantidos pelo governo brasileiro. De acordo com os dados que coletamos (veja Quadro 2), o Portal de Periódicos da CAPES e o Catálogo de Teses e Dissertações da CAPES possuem, cada um deles, cerca de 1,5 milhões de registros em seus bancos de dados. Trata-se de bancos de dados que se complementam – o primeiro agrega, majoritariamente, artigos de periódicos, enquanto o segundo armazena teses e dissertações. É lamentável que os sistemas de busca dos dois maiores portais científicos do Brasil tenham limitações que podem comprometer as revisões sistemáticas.

## 5 Conclusões

Ao realizar uma revisão sistemática, é importante utilizar variadas de fontes de informação, incluindo bancos de dados de artigos científicos, portais de teses e dissertações, dentre outras fontes relevantes para a pesquisa. Além disso, é importante seguir protocolos estabelecidos para a revisão sistemática, garantindo a transparência, replicabilidade e validade do processo de seleção e análise dos estudos incluídos.

Atualmente, os sistemas de busca de literatura científica online cobrem a maioria das disciplinas do campo do conhecimento e tornaram milhões de registros científicos acessíveis em segundos; em muitos casos, gratuitamente. O tempo que antigamente era despendido na procura física de evidências, agora é usado para o planejamento cuidadoso de pesquisas que utilizam sintaxe complexa e recursos de pesquisa para explorar dados textuais em títulos, resumos, palavras-chave e textos completos.

Dado que os sistemas de busca diferem em características técnicas, sua adequação para pesquisas sistemáticas pode variar. No entanto, os pesquisadores muitas vezes desconhecem as

características técnicas e as limitações dos sistemas de busca. Esse desconhecimento se dá, principalmente, entre alunos que estão começando a realizar suas primeiras buscas sistemáticas (BRINDESI; MONOPOLI; KAPIDAKIS, 2013). As avaliações descritas neste relatório demonstram que usar sistemas de busca acadêmica corretamente nem sempre é tão simples quanto as interfaces de busca podem sugerir.

A principal contribuição deste relatório foi realizar a avaliação de dez sistemas de busca, comumente usados por pesquisadores brasileiros, para realizar revisões sistemáticas na área de Ensino, segundo os critérios de Gusenbauer e Haddaway (2020). Avaliamos e descrevemos a usabilidade e a funcionalidade de sistemas de busca frequentemente utilizados em sínteses de evidências, com foco nas limitações que podem influenciar a qualidade das revisões sistemáticas.

Dos 10 sistemas de busca que avaliamos, cinco também foram avaliados por Gusenbauer e Haddaway (2020) – Pubmed, Scopus, Web of Science, Google Acadêmico e ERIC. Os demais são comumente utilizados por pesquisadores brasileiros, mas não são populares no exterior, e não foram avaliados pelos pesquisadores supracitados: Portal de Periódicos da CAPES, SciELO.br, BDTD, LILACS, Catálogo de Teses da CAPES. Os resultados da avaliação realizada pelos autores, em fevereiro e março de 2019, foram similares aos nossos. Isto é, na avaliação deles, ERIC<sup>3</sup>, PubMed, Scopus e Web of Science (*core collection*) também foram consideradas fontes primárias; enquanto Google Acadêmico foi considerada secundária.

Não testamos a integridade do conjunto de dados retornado pelos sistemas de busca. Este estudo examina o sistema de busca como uma ferramenta que faz a mediação entre uma base de dados de registros potencialmente relevantes e um pesquisador que deseja recuperar, analisar e sintetizar essa informação de uma forma sistemática e rigorosa.

Os testes mostraram diferenças substanciais entre os sistemas de busca. Embora alguns sistemas possam ser recomendados quase sem limitação, outros falharam em testes importantes, que limitam sua adequação para avaliações. Em outras palavras, nem todos os sistemas de busca permitem que os pesquisadores realizem consultas, apliquem filtros ou realizem pesquisa de citações com os altos padrões exigidos em revisões sistemáticas.

Nossa avaliação foi realizada em fevereiro e março de 2024 e pode sofrer alterações após essas datas, caso melhorias sejam feitas nos sistemas de busca. Além da avaliação dos dez sistemas, a contribuição desta pesquisa também está em descrever os critérios de Gusenbauer e Haddaway (2020) de uma maneira facilitada para pesquisadores brasileiros. Procuramos descrever os critérios para que sejam facilmente compreendidos, o que se faz importante principalmente para pesquisadores que ainda desconhecem as funcionalidades dos sistemas de busca acadêmica. Os critérios ajudam a tornar mais transparentes as qualidades e limitações dos sistemas de busca; com isso, esperamos também contribuir para que pesquisadores experientes de todas as áreas possam realizar avaliações nos sistemas de busca que julgarem importantes para suas pesquisas.

---

<sup>3</sup>A base de dados do ERIC é acessível através do seu sistema de pesquisa dedicado, mas também através do EbscoHost. Em nossos testes, utilizamos o sistema dedicado, disponível em <https://eric.ed.gov/>. Gusenbauer e Haddaway (2020) utilizaram EbscoHost, então alguns resultados podem ser diferentes.

## Referências

BRAMER, W. et al. The comparative recall of Google Scholar versus PubMed in identical searches for biomedical systematic reviews: A review of searches used in systematic reviews. *Systematic reviews*, v. 2, p. 115, dez. 2013.

BRINDESI, H.; MONOPOLI, M.; KAPIDAKIS, S. Information Seeking and Searching Habits of Greek Physicists and Astronomers: A Case Study of Undergraduate Students. *Procedia - Social and Behavioral Sciences*, v. 73, p. 785–793, fev. 2013. ISSN 1877-0428. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877042813004126>>.

CAMPBELL. *The Campbell Collaboration*. 2025. Disponível em: <<https://www.campbellcollaboration.org/>>.

CEE. *Environmental Evidence – Reliable evidence, informed decisions, better environment*. 2025. Disponível em: <<https://environmentalevidence.org/>>.

COCHRANE. *Cochrane - Trusted evidence. Informed decisions. Better health*. 2025. Disponível em: <<https://www.cochrane.org/>>.

GUSENBAUER, M.; HADDAWAY, N. R. Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research Synthesis Methods*, v. 11, n. 2, p. 181–217, mar. 2020. ISSN 1759-2879, 1759-2887. Disponível em: <<https://onlinelibrary.wiley.com/doi/10.1002/jrsm.1378>>.

HADDAWAY, N. R. et al. The Role of Google Scholar in Evidence Reviews and Its Applicability to Grey Literature Searching. *PLOS ONE*, v. 10, n. 9, p. e0138237, set. 2015. ISSN 1932-6203. Publisher: Public Library of Science. Disponível em: <<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0138237>>.

HJØRLAND, B. Classical databases and knowledge organization: A case for boolean retrieval and human decision-making during searches. *Journal of the Association for Information Science and Technology*, v. 66, n. 8, p. 1559–1575, 2015. ISSN 2330-1643. \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.23250>. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.23250>>.

KITCHENHAM, B.; CHARTERS, S. *Guidelines for performing systematic literature reviews in software engineering*. Technical report, EBSE Technical Report EBSE-2007-01, 2007. Disponível em: <<https://docs.edtechhub.org/lib/EDAG684W>>.

PRISMA. *PRISMA statement*. 2020. Disponível em: <<https://www.prisma-statement.org/>>.

ROSES. *Systematic Review Reporting Standards*. 2025. Disponível em: <<https://www.roses-reporting.com/>>.