

# Phrasing in Reproducible Search Methodology: The Consequences of Straight and Curly Quotation Marks

Katie Barrick and Amy Riegelman

In recent years, various disciplines have engaged in efforts to increase research reproducibility including the adoption of replicable search methodologies. With the development of reporting checklists and guidelines for systematic reviews such as the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Statement, authors are expected to transparently report search strategies. Replicable search strategies are critical since the included studies will be screened for inclusion in some forms of evidence synthesis, which could have practice and policy implications. In cases where search strategies miss germane literature, studies are open to further criticism and can face difficulties in peer review. In some cases, search strategies that contain nonalphanumeric or special characters may not retrieve pertinent literature due to a search platform's capacity for handling said characters. In this study, we explore issues with phrasing search strategies containing curly and straight quotation marks tested in 40 search platforms. The discovery platforms were tested using quotation characters and the absence of said characters to investigate platform behavior. Searches were categorized into one of five groups: CI (curly ignored), CA (curly acknowledged), CU (curly unclear), NP (no phrases), and UC (unsupported characters). The study found that 42.5 percent of platforms ignored curly quotation marks and interpreted the test term as a phrase, 30 percent of the bibliographic platforms acknowledged curly quotations and completed the phrase search, and one platform flagged curly quotation marks as an unsupported character.

## Introduction

In this article we examine search platform functionality regarding double quotation marks. Due to reporting checklists and guidelines for systematic reviews and other evidence synthesis methods (such as meta-analyses and scoping reviews), there is an expectation that authors report key details about their search. Search strategies and the search syntax affect the reproducibility of the search. In a perfect world, researchers should be able to perform a search as transparently reported in a manuscript and receive the same results used in the sample of a review. In evidence synthesis, the studies that are screened and coded are the crux of the research. If pertinent studies are not

---

\*Katie Barrick is Biosciences Librarian and Data Curator in Magrath Library at the University of Minnesota; email: wilsonkm@umn.edu. Amy Riegelman is Social Sciences Librarian in Wilson Library at the University of Minnesota; email: aspringe@umn.edu. ©2021 Katie Barrick and John Raible, Attribution-NonCommercial (<https://creativecommons.org/licenses/by-nc/4.0/>) CC BY-NC.

included, authors may face resistance as a manuscript is peer-reviewed or publicly criticized via social media. Examples of criticized reviews include a Cochrane systematic review critiqued for not including several eligible randomized trials and an entomology meta-analysis claiming decline of the entomofauna, which has been coined as insectageddon.<sup>1</sup> In the latter case, the researchers used the keyword *declin\** in a Web of Science (WoS) search that skewed the studies included in their sample. As acknowledged by many scientists including Thomas, Jones, and Hartley, using *declin\** “immediately biases the meta-analysis toward exaggerated estimates of decline rates, even assuming there is no underlying publication bias in the literature.”<sup>2</sup>

The present study is relevant to the work of academic libraries because the librarian skillset is valued in evidence synthesis methodologies. The librarian role and value in supporting systematic review literature searches is well documented in Campbell and Cochrane Handbooks as well as other research networks.<sup>3</sup> Additionally, studies conducted by librarians have shown that librarian-involved published reviews are more reproducible.<sup>4</sup> The work of librarians to transparently report the search methods is necessary for either reproducing a search or replicating and extending a review. Without being provided the key details of information retrieval methods, a replication attempt is very challenging, if not impossible, as has been shown in attempted replications of empirical cancer research in Reproducibility Project: Cancer Biology.<sup>5</sup> Without transparently reported research methods and/or cooperation from original authors, other scholars struggle to replicate.<sup>6</sup>

For the purposes of this paper, we specifically studied phrasing strategies wherein curly or straight quotation marks are used. *Straight* is used to describe quotation marks (sometimes called the double quote) labeled below in table 1. Alternative names for straight quotation marks are neutral, vertical, dumb, typewriter, or American Standard Code for Information Interchange (ASCII) quotation marks. Left and right straight quotation marks are identical and both symmetrical. Alternative names for *curly* quotation marks include typographic, smart, curved, or book.

**TABLE 1**  
**Depiction and Character Codes of Straight and Curly Quotation Marks**

	Curly Quotation Marks		Straight Quotation Marks	
	Left	Right	Left	Right
	"	"	"	"
ASCII Character Code	NA	NA	00100010	00100010
Unicode Character Code	U+201C	U+201D	U+0022	U+0022

For this study, readers may benefit from familiarity of both ASCII and Unicode, character encoding systems. ASCII has 128 characters, was originally developed in 1963, and evolved from telegraph and teleprinter computing.<sup>7</sup> The 128 characters are represented by seven-digit binary sequences of zeros and ones. The binary sequence for quotation marks is 00100010.<sup>8</sup> ASCII’s limit of 128 characters is due to the 128 combinations of seven digits of zeros and ones.<sup>9</sup>

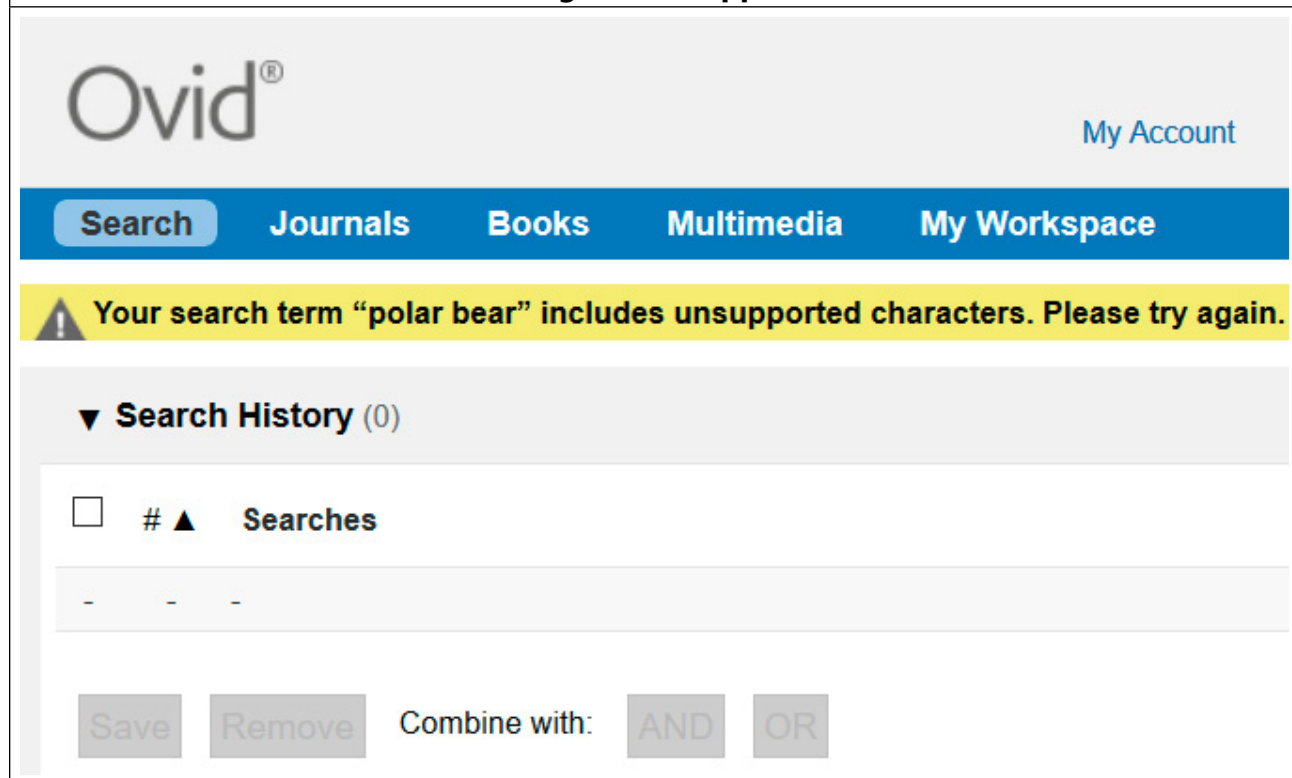
Developed in 1991, Unicode is more compatible with international writing systems and languages. Characters are assigned unique identification numbers, also known as code points. Computer operating systems then use the code points to identify characters within font files, resulting in how the characters are displayed. A dramatic difference between ASCII and Unicode is in the number of available characters. Whereas ASCII has only 128 characters, there are 137,000 Unicode characters as of May.<sup>10</sup>

Curly quotation marks do not exist in the ASCII character set. For the purposes of this study, it is important to recognize that, if curly quotation marks are used in a bibliographic platform that is not compatible with curly quotation marks, the search results could reflect a number of different scenarios as reflected in the results section of this manuscript.

When scholars seek peer-reviewed and grey literature on a given topic, they have many options for search platforms, and these platforms have differing search syntaxes and algorithms for displaying results that can shift over time without any indication to the user. One common search functionality is phrasing; in phrase searching, a searcher is able to seek results with the exact words in the exact order. When a searcher inputs a search phrase (such as “polar bear”), they perhaps expect the search results to list only results that contain those words in the exact order. Before engaging in this study, the authors were aware of at least one search platform, Ovid, that informs searchers that curly quotation marks are an unsupported character (see figure 1). Among other search platforms, this messaging is inconsistent or nonexistent. Assumedly, misconceptions about search input and special characters are partly due to the “blackbox” or proprietary nature of search products and the absence of technical knowledge needed to fully understand search engines.

Complicating matters, the direct input of quotation marks in Google Docs and Microsoft Word (specifically Office 365, Office Online, Office 2019, Office 2016, Office 2013, Office 2010) automatically appear as curly quotation marks. If a search strategy is developed in one of these word processors and copied/pasted into a search platform, the curly quotation marks could become problematic in that the phrasing is completely ignored or the search is flagged as having unrecognized characters.

**FIGURE 1**  
**Ovid Error Message for Unsupported Characters**



Phrase search is particularly important for systematic reviews. Systematic reviews aim to collect and analyze evidence regarding a specific research question and to provide a summary of said evidence. According to guidance from the Center for Disease Control, literature searching for systematic review purposes should consider both recall and precision—“Recall (sensitivity) is defined as the number of relevant reports identified divided by the total number of relevant reports in existence. Precision (specificity) is defined as the number of relevant reports identified divided by the total number of reports identified.”<sup>11</sup>

Some research methodologies require transparent reporting of search strategies. Examples include systematic reviews, meta-analyses and other evidence synthesis methodologies compliant with reporting guidelines and checklists such as PRISMA. These documents include guidance on what search method details need to be reported in a manuscript or supplemental materials. The guidelines often mention the need to report search strategies and search syntaxes to ensure that the search strings could be input by another user and therefore result in the search query producing the same results.

### ***PRISMA Statement, Checklist, and Extensions***

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) is intended to help authors ensure transparent and reproducible methods. In 2009 PRISMA replaced the pre-existing Quality of Reporting of Meta-analyses (QUOROM). As of January 2020, 185 journals and editorial organizations have endorsed PRISMA.<sup>12</sup> Adhering to PRISMA requires that authors transparently report a full search strategy for at least one database and include any limits.<sup>13</sup> The transparent reporting of this search strategy is often reported in the appendix of a manuscript. Depending on the search itself or stylistic preferences of the journal, the published search strategy could include either straight or curly quotation marks and may be changed from the authors’ manuscript during copy editing. PRISMA indicates that the transparent search be communicated in a way that allows for it to be reproducible. This may mean that a search strategy is copied and pasted into a search platform to reproduce or replicate (such as by extension or update) the reported search methods. Examples below identify language used in PRISMA documentation relevant to the present study.

#### **PRISMA Checklist:**

“Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.”

#### **PRISMA for Scoping Reviews (PRISMA-ScR)**

Item 8:

“Report the search in a manner that allows for easy replication by others” and “Provide the entire search strategy for at least one database either in the text, a table or an Appendix”

**Preferred Reporting Items for Systematic review and Meta-Analysis Searches (PRISMA-S)** provides guidance for expert searching and reporting. The authors explain the need for this extension: “Each platform offers different ways of searching the databases, such as platform-specific *field codes*..., phrase searching, truncation, or searching full-text versus abstract and keyword only. Different platforms may contain additional data that are not available in the original database, such as times cited, social media impact, or additional

keywords. These differences between the platforms can have a meaningful impact on the results provided.”<sup>14</sup> This extension advises that searchers report the “search strategies for each database and information source, copied and pasted exactly as run.”<sup>15</sup>

*Methodological Expectations of Campbell Collaboration Intervention Reviews: Reporting standards (MECCIR)* are the standards for Campbell Collaboration reviews. These reporting standards indicate that reviews are required to present the search strategies “used for each database in an Appendix, including any limits and filters used, so that it could be replicated.”<sup>16</sup>

*Methodological Expectations of Cochrane Intervention Reviews (MECIR)* are the standards for Cochrane reviews. C36 of the *Cochrane Handbook for Systematic Reviews of Interventions* indicates that the search strategy must be reported with enough detail to ensure that “all the searches of all the databases are reproducible.”<sup>17</sup>

#### *Meta-analysis of observational studies in epidemiology (MOOSE)*

These standards are specifically for meta-analyses of observational studies in epidemiology. This document explains that the search methods should include the search strategy including keywords, databases, and “software used, name and version, including special features used (e.g., explosion).”<sup>18</sup>

#### *Peer Review of Electronic Search Strategies (PRESS)*

PRESS is a tool used to peer-review search strategies. This source reinforces that accurate reporting is needed “to ensure critical appraisal, replication, and updating.”<sup>19</sup> This tool prompts users to consider ways in which precision could be improved, and this text includes language on phrasing: “Could precision be improved by using proximity operators (e.g., adjacent, near, within) or phrase-searching instead of AND?”<sup>20</sup> Users are asked to consider search syntax differences between search platforms. Specifically: “Are there any errors in system syntax; for example, the use of a truncation symbol from a different search interface?”<sup>21</sup>

## Literature Review

While the reporting guidance offered by guidelines and checklists is valuable for consistency across various studies and potentially improved reproducibility and replicability, an exploration into the scholarly literature on special characters revealed two different themes: concern over the use of special or nonalphanumeric characters in literature and the perceived impact of these characters on citation counts.

The concern regarding the use of special characters or nonalphanumeric characters in scholarly literature and their perceived impacts on citation counts of articles may be due to issues with search retrieval. Hartley asserts the citation counts of article titles containing colons, also known as titular colonicity, are not negatively impacted by the special character, while Jamali and Nikazad reported titular colonicity does negatively impact citation counts.<sup>22</sup>

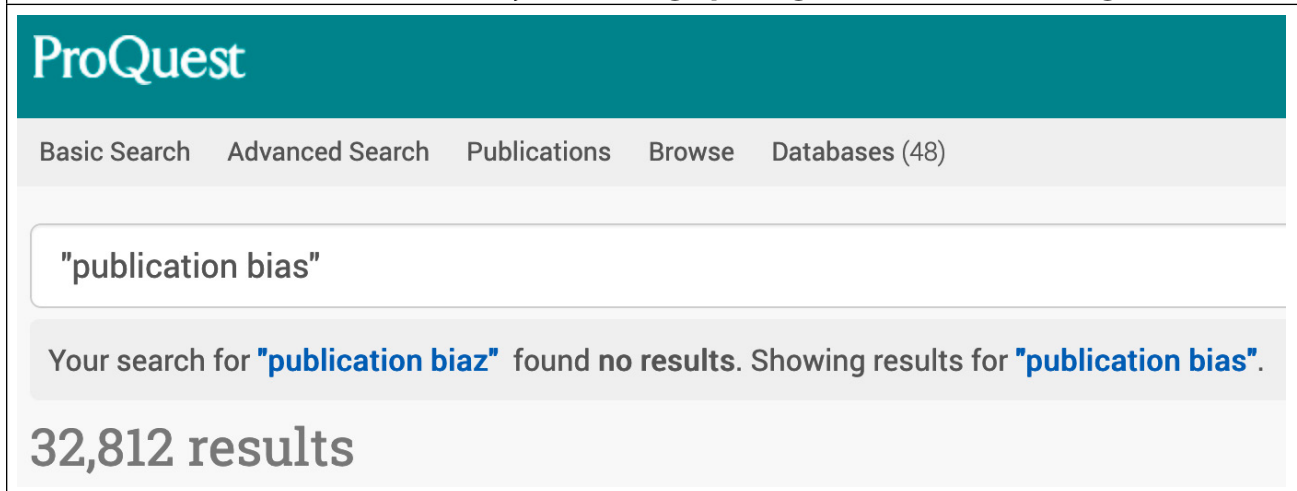
The use of nonalphanumeric characters in article titles is increasing, from colons to question marks.<sup>23</sup> Ball reported the number of life sciences articles using question marks in titles increased from 50 to 200 percent over four decades, from 1966 to 2005.<sup>24</sup> Similarly, Buter and

Raan examined the use of 29 nonalphanumeric characters in the titles of scientific articles and found 68 percent of the sampled WoS articles, published between 1999 and 2008, contained at least one nonalphanumeric character.<sup>25</sup>

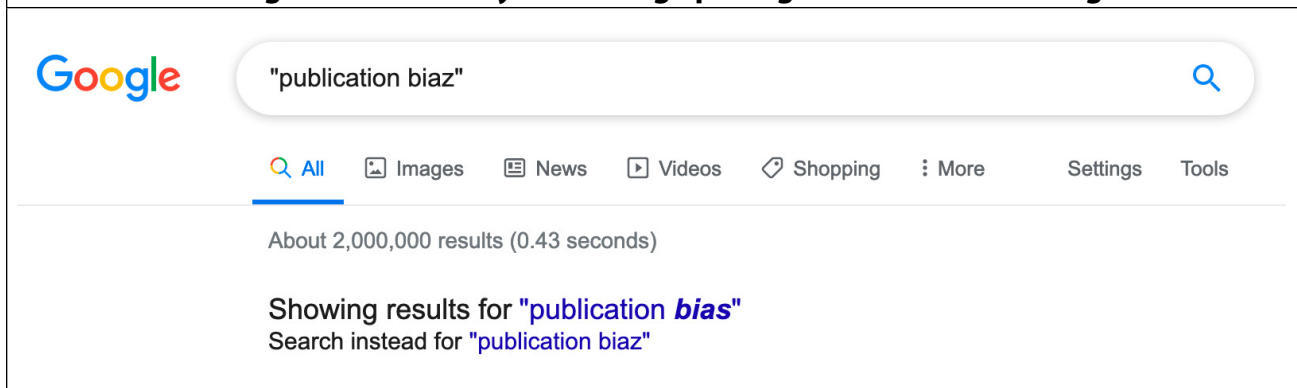
The increase of nonalphanumeric characters in scholarly literature paired with developments in citation analysis research are impacted by an additional factor: the capacities of bibliographic platforms. Zhou, Tse, and Witheridge examined the robustness of Scopus and WoS using metamorphic relations and found articles with titles containing hyphens negatively impacted citation counts.<sup>26</sup> The results of the Zhou et al. article “[challenged the common belief that citation counts and journal impact factors (JIFs)] are reliable measures of the impact of papers and journals, as they can be distorted simply by the presence of hyphens in paper titles.”<sup>27</sup> WoS issued a response to Zhou et al. stating that article titles are not used to link citation to source records as they are “entirely aware that this field can be ambiguous and prone to error” and claimed that hyphens within article titles do not prevent the linking between a source and a cited reference.<sup>28</sup> WoS also analyzed unlinked citations wherein titles contained hyphens and found issues were actually linked to citation errors and/or missing metadata.

Studies on phrase searching also identified limitations and errors. In a study conducted by Salvador-Oliván, Marco-Cuenca, and Arquero-Avilés about the impact of errors in search strategies on the quality and validity of systematic reviews, 5.8 percent of their sample included

**FIGURE 2**  
**ProQuest Automatically Correcting Spelling Error in Search String**



**FIGURE 3**  
**Google Automatically Correcting Spelling Error in Search String**



errors in searching for phrases.<sup>29</sup> An example of the challenges of search platform differences revolves around search truncation in PubMed, wherein only the last term in a phrase should be truncated to retrieve items correctly. Salvador-Oliván et al. note that “it is also important to be knowledgeable regarding the principles of information retrieval in order to avoid committing basic errors and to apply these principles to the particular characteristics of the search language of the database used.”<sup>30</sup> A gap in this type of knowledge creates the kind of concern identified in literature regarding special characters or nonalphanumeric characters in literature.

**TABLE 2**  
**Search Syntax Terms and Definitions**

<b>Terms</b>	<b>Definitions</b>
input	word or words (in the form of characters) inserted into a platform’s search query area
search query	word or words (in the form of characters) used to request information from a search engine
search engine results page (SERP)	display of results in an order according to an algorithm in response to query
recall (sensitivity)	“Recall (sensitivity) is defined as the number of relevant reports identified divided by the total number of relevant reports in existence.” (CDC)
precision (specificity)	“Precision (specificity) is defined as the number of relevant reports identified divided by the total number of reports identified.” (CDC)
query expansion	search platform attempts to improve information retrieval performance by expanding a query with one or more techniques (a couple of examples: stemming, correcting spelling errors)
lemmatization	use of an algorithm to determine the inflected words. If the base word (mouse) is used in a query, inflected forms (mice) could influence the search results
stemming	Regardless of context, the base form of terms treat other terms as synonymous if they have a common base word (for example: race, races, racecar). When a search engine stems words, they are referred to as stemmers, stemming programs, or stemming algorithms
validation	input is checked to ensure certain criteria (such as unclosed quotation marks)
sanitization	input is modified to meet validation criteria (like adding the missing quotation marks)
search syntax	could vary by search platform but generally includes a search string with fielded (examples: title, author) searches, compound queries with boolean operators (AND, OR, NOT), or phrasing
character encoding	allows a computer to interpret numbers as representing characters; a search word or words includes characters from a character set
phrase-searching	query is performed to reflect search terms in an exact order; example of precision
search truncation	a portion of a word is searched, usually with a symbol such as an asterisk, to specify the beginning of a term but retrieving variations on the term (for example, <i>search*</i> retrieves sources containing the terms <i>search</i> , <i>searcher</i> , <i>searchers</i> , <i>searches</i> , and <i>searching</i> )
metamorphic testing	metamorphic testing is a software technique used to address the test oracle problem, wherein the oracle is a mechanism used to determine whether or not test outputs match expected test outcomes
metamorphic relations	relations between inputs and outputs found after metamorphic testing

Additionally, platforms have the ability to change how phrase search works when attempted by users. This ability is at times more noticeable to users such as query expansion. For example, both Google and ProQuest will correct a misspelled query such as “publication biaz” to “publication bias” and inform users of the change on the search results page.

In other cases, changes may not be as explicit or clearly communicated to users, such as updates made in ProQuest’s search interface in 2010.<sup>31</sup> As documented by Notess, the legacy version of ProQuest would search queries with two terms (for instance: publication bias) as a phrase search by default, but the newer version made changes to how multiple term queries were performed.<sup>32</sup> The newer version assumes an AND operator exists between each individual term. ProQuest is just one example of a vast array of search interfaces with the ability to change syntaxes with or without any notification to the users. This reduces the transparency of the search algorithms and, by extension, may reduce the reproducibility of the study.

## Methods

The purpose of this study is to examine if straight quotation marks and curly quotation marks perform differently in various search platforms. The methods for this study were preregistered on the Open Science Framework on March 27, 2019.<sup>33</sup> The authors compared the total search results in each platform for curly, straight, and no quotation marks. We collected data for each platform and then analyzed how the platforms responded to curly and straight quotation marks in an exact phrase search. We expected to observe one of the following outcomes for each platform:

1. The search results when using straight quotation marks and curly quotation marks were the exact same
2. There was a variance between search results when using straight quotation marks compared to using curly quotation marks
3. The use of curly quotation marks was rejected due to unsupported characters

Forty platforms were tested between March 28, 2019 and April 8, 2019. These dates were captured as search performance may vary greatly due to changes within search platforms. The sample was selected based off institutional usage statistics, grey literature platforms desirable for systematic reviews, and uniqueness. Uniqueness was part of the selection criteria because some platforms operated under the same search syntax. For example, Ovid Medline and PsycINFO via Ovid are expected to react the same to straight and curly quotation marks. The platforms selected represent indexing of literature from many different disciplines.

The following platforms were tested by the authors, and result totals were input into a spreadsheet. The full list of platforms tested is viewable in table 3. Variations of *publication bias* were used as a search query and tested three times in each platform: once using straight quotation marks, once using curly quotation marks, and once without quotation marks.

<b>Observed Platform</b>	<b>Results Straight</b>	<b>Results Curly</b>	<b>Results Without Quotes</b>	<b>Results Performance</b>
AGRIS	16	41	41	CI
EThOS (e-theses online service)	39	384	384	CI
F1000Research	57	532	532	CI



**TABLE 3**  
**Search Platform Results According to Quotation Mark Variations**

<b>Observed Platform</b>	<b>Results Straight</b>	<b>Results Curly</b>	<b>Results Without Quotes</b>	<b>Results Performance</b>
Gale Databases	8	16	16	CI
HathiTrust	691	839,014	839,014	CI
Hispanic American Periodicals Index (HAPI)	0	1	1	CI
OpenGrey	14	23	23	CI
OSF Preprints	43	401	401	CI
OSTI	53	384,758	384,758	CI
PeerJ Preprints	173	3,506	3,506	CI
POPLINE	71	124	124	CI
SHARE	2,461	397,533	397,533	CI
Social Science Open Access Repository	6	76	76	CI
SSRN	160	268	268	CI
Project Muse	82	6,341	6,341	CI
PubAg	403	542	542	CI
WHO Library and Information Networks for Knowledge	370	5,928	5,928	CI
arXIV	22	226	1,721	CU
Campbell Collaboration Library	19	0	32	CU
IDEAS (RePEc)	201	0	3,356	CU
INSPIRE High-Energy Physics	5	0	1,457	CU
OCLC (like WorldCat, OALster)	2263	0	13,182	CU
Science.gov	1,429	2,208	2,411	CU
bioRxiv	187	187	5,598	CA
Cochrane Library	217	217	1,111	CA
JSTOR	3,126	3,126	101,838	CA
ERIC via Department of Education	96	96	877	CA
EBSCO Databases	24,260	24,620	25,817	CA
Google Scholar	324,000	324,000	3,950,000	CA
HeinOnline	541	541	201,100	CA
PubMed	13,716	13,716	16,111	CA
Scopus	14,810	14,810	25,736	CA
Proquest	196	196	25,639	CA
Theses Canada	16	16	43	CA
Web of Science	12,706	12,706	18,618	CA
Mathematical Physics Preprint Archive	12	12	12	NP
NCJRS	35	35	35	NP
Columbia International Affairs Online (CIAO)	1	1	1	NP
Preprints.org	10	10	10	NP
Ovid	13,663	0	13,663	UC

## Results

Results were organized into five categories: CI (curly ignored), CU (curly unclear), CA (curly acknowledged), NP (no phrases), and UC (unsupported characters). Table 3 reflects the search results and assigned category.

Of the 40 bibliographic platforms tested, 17 (42.5%) platforms ignored curly quotation marks (CI) and did not interpret the test term as a phrase. Four (10%) platforms did not recognize either straight or curly quotation marks and did not recognize phrase search. One platform flagged curly quotation marks as an unsupported character. Only 12 (30%) of the bibliographic platforms acknowledged curly quotations and completed the phrase search when used. For the remaining six (15%) platforms, curly quotation marks were not treated the same as straight quotation marks, but it was unclear to the authors as to how they were read and interpreted.

The six platforms categorized as “CU” were contacted to clarify platform functionality with respect to curly quotations. Five out of the six platforms responded. Four organizations affirmed it was unclear how the search engine read and interpreted curly quotation marks and passed our query on to platform developers. The remaining organization shared that straight quotation marks were recognized, while curly quotation marks were ignored, thus resulting in terms being searched independently.

## Discussion

As indicated by the results, there are consequences to using either straight or curly quotation marks in the various search interfaces. The authors have recommendations based on these findings.

### *Recommendations*

In order for search strategies to be transparently reported and reproducible, search platforms should transparently report the search syntax as well as when or if modifications are made to the search syntax. Since search strategies abiding by evidence synthesis guidelines are required to report the date of the search, a syntax change could affect the number of search results from one day to the next. When search platforms act as a blackbox, searchers and potentially evidence-based practice and policy are at a disadvantage. Search platforms could transparently report changes similar to how software companies maintain release notes or change logs. Release notes are used to communicate product updates such as changes in functions, deprecations, or entirely new features. They are typically versioned as either major releases (such as a new feature) or minor releases (like a minor bug fix).

Librarians, information specialists, and other expert searchers performing comprehensive and transparently reported searches need to be aware of how a search query is or is not affected by different characters. Studies like this could bring more awareness to this topic. Authors and copy editors need to maintain the integrity of the search strategy as performed. For example, if a search string is reported as “publication bias,” the straight quotation marks should not be modified to curly quotation marks due to journal stylistic preferences. The straight quotation marks in this instance could serve an important purpose. Changing the quotation marks from straight to curly could alter the search reproducibility.

Settings on both Google Docs and Microsoft Word could be adjusted to default to straight quotation marks. To change settings on Google Docs, select *tools, preferences* and then uncheck

*Use Smart Quotes.* For Microsoft Word 2016, the settings could be adjusted by selecting *Files, Options,* and then *Proofing* followed by *AutoCorrect.* Then in the *Replace as you type* section, input “ and “ in the *Replace* category followed by " and " in the *With* category. Specific to reproducible and transparent literature searching, one would first need to know that curly quotations could cause oddities in the search results, and following, they would need to take several additional steps to revert the quotation marks to straight.

### *Journal Stylistic Preferences*

The issue is exacerbated by stylistic preferences of journals committed to style guides. Ten style guides representing different disciplines were consulted for their stance on quotation marks. While all style guides contained some direction regarding quotation marks, guidance was generally limited to usage and proper placement within the body of text. Of the 10 style guides, one guide gave clear direction on the type of quotations to use: *The Chicago Manual of Style.* *The Chicago Manual of Style* states the preference for curly quotation marks and points.

“Published works should use directional (or ‘smart’) quotation marks, sometimes called typographer’s or ‘curly’ quotation marks. These marks, which are available in any modern word processor, generally match the surrounding typeface. For a variety of reasons, including the limitations of typewriter-based keyboards and of certain software programs, these marks are often rendered incorrectly. Care must be taken that the proper mark—left or right, as the case may be—has been used in each instance.”<sup>34</sup>

The conflict between stylistic preferences and reproducible search strategies may have consequences. Should a journal alter a search strategy that contains straight quotation marks

**TABLE 4**  
**Style Manuals and Curly/Straight Quotation Mark Preferences**

<b>Style Guide</b>	<b>Version Referenced</b>	<b>Curly/Straight Preference</b>	<b>Preference Guidance</b>
The ACS Style Guide: Effective Communication of Scientific Information	3rd	None	N/A
American Medical Association	10th	None	N/A
American Psychological Association	6th	None	N/A
American Sociological Association	5th	None	N/A
Bluebook	20th	None	N/A
Chicago Manual Style	17th	Curly	CMOS Online, 6.115: “Smart” quotation marks
Council of Science Editors	8th	None	N/A
IEEE	V 11.12.18	None	N/A
Modern Language Association	8th	None	N/A
Vancouver/NLM’s Citing Medicine	2nd	None	N/A

to curly quotation marks, the stylistic choice would therefore alter the reproducibility of the search. Despite the fact that many style manuals do not refer to quotation marks (see table 4), curly quotation marks are generally a desired stylistic choice made by journals.

### ***Limitations***

Our study is a snapshot in time. The dates that the searches were performed matters greatly in terms of our results due to the ability of a search platform to make significant search syntax changes and not inform searchers. The platforms may be revisited in the future to investigate how previously retrieved literature may change in the interim due to unadvertised changes to search syntaxes. The characters preferred in journals evolve over time. More could be explored regarding direct search input of various devices—that is, iOS. Stylistically, manuals change periodically as well. Our study is also focused on characters belonging to the English language and is centered on the United States. Additional research needs to be done with other languages and more inclusive of a global perspective.

Future research should study the vast array of characters as well as search platform transparency regarding search syntax. Examples of characters that should be studied include emoticons and emojis, which have started to appear in newspaper article titles. For example, the shruggy emoticon, “\\_(ツ)\_/”, appears in 11 article titles in ProQuest’s U.S. Newsstream as of January 22, 2021.

### ***Implications for Higher Education and Libraries***

As indicated earlier, a search platform’s capacity for handling special and nonalphanumeric characters may impact the replicability of search strategies used in systematic reviews and meta-analyses conducted by researchers and librarians. The lack of clarity regarding changes made to search syntax by platforms may further complicate the trustworthiness of and replicability of some studies, as researchers and librarians may not be aware of said changes. Further education, communication, and awareness must be brought regarding the robustness of search platform capacities and specialized search.

---

## **Notes**

1. L. Jørgensen, P.C. Gøtzsche, and T. Jefferson, “The Cochrane HPV Vaccine Review Was Incomplete and Ignored Important Evidence of Bias,” *BMJ Evidence-Based Medicine* 23, no. 5 (2018): 165–68; C.D. Thomas, T.H. Jones, and S.E. Hartley, “Insectageddon: A Call for More Robust Data and Rigorous Analyses,” *Global Change Biology* 25, no. 6 (2019): 1891–92.

2. Thomas, Jones, and Hartley, 1891–92.

3. Kugley, S., Wade, A., Thomas, J., Mahood, Q., Jørgensen, A. M. K., Hammerstrøm, K., & Sathe, N. (2016). “Searching for studies: A guide to information retrieval for Campbell,” *Campbell Systematic Reviews*, [https://campbellcollaboration.org/images/Campbell\\_Methods\\_Guides\\_Information\\_Retrieval.pdf](https://campbellcollaboration.org/images/Campbell_Methods_Guides_Information_Retrieval.pdf); *Cochrane Handbook for Systematic Reviews of Interventions*, v. 6.1 (updated September 2020), eds. J.P.T Higgins et al. (London, UK: Cochrane, 2020), available from [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook); Genevieve C. Gore and Julie Jones, “Systematic Reviews and Librarians: A Primer for Managers,” *Partnership: The Canadian Journal of Library and Information Practice and Research* 10, no. 1 (2015).

4. Jonathan B. Koffel, “Use of Recommended Search Strategies in Systematic Reviews and the Impact of Librarian Involvement: A Cross-sectional Survey of Recent Authors,” *PloS One* 10, no. 5 (2015): e0125931; Jonathan B. Koffel and Melissa L. Rethlefsen, “Reproducibility of Search Strategies Is Poor in Systematic Reviews Published in High-impact Pediatrics, Cardiology and Surgery Journals: A Cross-sectional Study,” *PLoS One* 11, no. 9 (2016): e0163309.

5. Timothy M. Errington et al., “Science Forum: An Open Investigation of the Reproducibility of Cancer Biology Research,” *Elife* 3 (2014): e04333.

6. Errington, et al, e04333.
7. ASCII (May 16, 2019), *Encyclopædia Britannica*, <https://www.britannica.com/topic/ASCII> [accessed 2 December 2019].
8. ASCII, *Encyclopædia Britannica*.
9. ASCII, *Encyclopædia Britannica*.
10. Unicode (November 14, 2019), *Encyclopædia Britannica*, <https://www.britannica.com/topic/Unicode> [accessed 2 December 2019].
11. Systematic Reviews (n.d.), *Centers for Disease Control and Prevention*, retrieved from <https://www.cdc.gov/library/researchguides/systematicreviews.html>.
12. PRISMA Endorsers, *PRISMA*, retrieved from [www.prisma-statement.org/Endorsement/PRISMAEndorsers](http://www.prisma-statement.org/Endorsement/PRISMAEndorsers).
13. PRISMA Statement, *PRISMA*, retrieved from [www.prisma-statement.org/PRISMAStatement/](http://www.prisma-statement.org/PRISMAStatement/).
14. M.L. Rethlefsen et al., and Prisma-S. Group, "Prisma-S: An Extension to the Prisma Statement for Reporting Literature Searches in Systematic Reviews," *Systematic Review* 10, no. 1 (January 26, 2021): 39.
15. Rethlefsen et al., "Prisma-S."
16. "Methodological Expectations of Campbell Collaboration Intervention Reviews: Reporting Standards (MECCIR)" (October 2019), *Campbell Policies and Guidelines Series No. 4*.
17. Higgins et al., *Cochrane Handbook*, chapter 4.
18. D.F. Stroup et al., "Meta-analysis of Observational Studies in Epidemiology: A Proposal for Reporting," *JAMA* 283, no. 15 (2000): 2008–12, <https://doi.org/10.1001/jama.283.15.2008>.
19. J. McGowan et al., "PRESS Peer Review of Electronic Search Strategies: 2015 Guideline Statement," *Journal of Clinical Epidemiology* 75 (2016): 40–46, <https://doi.org/10.1016/j.jclinepi.2016.01.021>.
20. McGowan et al., 40–46.
21. McGowan et al., 40–46.
22. James Hartley, "Planning That Title: Practices and Preferences for Titles with Colons in Academic Articles," *Library & Information Science Research* 29, no. 4 (2007): 553–68, <https://doi.org/10.1016/j.lisr.2007.05.002>; Hamid R. Jamali and Mahsa Nikzad, "Article Title Type and Its Relation with the Number of Downloads and Citations," *Scientometrics* 88, no. 2 (2011): 653–61, <https://doi.org/10.1007/s11192-011-0412-z>.
23. J.T. Dillon, "In Pursuit of the Colon," *Journal of Higher Education* 53, no. 1 (1982): 93–99, <https://doi.org/10.1080/00221546.1982.11780427>; Rafael Ball, "Scholarly Communication in Transition: The Use of Question Marks in the Titles of Scientific Articles in Medicine, Life Sciences and Physics 1966–2005," *Scientometrics* 79, no. 3 (2009): 667–79, <https://doi.org/10.1007/s11192-007-1984-5>.
24. Ball, "Scholarly Communication in Transition."
25. R.K. Buter and A.F.J. van Raan, "Non-Alphanumeric Characters in Titles of Scientific Publications: An Analysis of Their Occurrence and Correlation with Citation Impact," *Journal of Informetrics* 5, no. 4 (2011): 608–17, <https://doi.org/10.1016/j.joi.2011.05.008>.
26. Zhi Quan Zhou, T.H. Tse, and Matt Witheridge, "Metamorphic Robustness Testing: Exposing Hidden Defects in Citation Statistics and Journal Impact Factors," *IEEE Transactions on Software Engineering* (2019): 1, <https://doi.org/10.1109/tse.2019.2915065>.
27. Zhou, Tse, and Witheridge, "Metamorphic Robustness Testing," 1.
28. Web of Science, "Web of Science Group Response to Metamorphic Robustness Testing," Clarivate.com, <https://clarivate.com/webofsciencegroup/blog/web-of-science-group-response-to-metamorphic-robustness-testing-exposing-hidden-defects-in-citation-statistics-and-journal-impact-factors/> [accessed 26 June 2020].
29. José Antonio Salvador-Oliván, Gonzalo Marco-Cuenca, and Rosario Arquero-Avilés, "Errors in Search Strategies Used in Systematic Reviews and Their Effects on Information Retrieval," *Journal of the Medical Library Association* 107, no. 2 (2019), <https://doi.org/10.5195/jmla.2019.567>.
30. Salvador-Oliván, Marco-Cuenca, and Arquero-Avilés, "Errors in Search Strategies Used in Systematic Reviews and Their Effects on Information Retrieval."
31. ProQuest Support, online chat with author, January 8, 2020.
32. Greg R. Notess, "Intricacies of Phrase Searching," *Online Searcher* 38, no. 6 (2014): 65–67.
33. Amy Riegelman and Katie Barrick, "Phrasing in Reproducible Search Methodology: The Quirks of Straight and Curly Quotation Marks," *OSF Registry* (March 27, 2019), [osf.io/qm3g7](https://osf.io/qm3g7).
34. *The Chicago Manual of Style*, 17th ed. (Chicago, IL: University of Chicago Press, 2017).