# Ethics and Practice of Knowledge Integrity in Communicating Health and Medical Research

Scott A. Mogull

## Abstract

Rhetoric of health and medicine (RHM) knowledge integrity is explored in the context of preparing RHM students, researchers, and practitioners to be careful curators and communicators of information from the medical literature. More specifically, the goal of this article is to provide a systematic framework for researching and citing claims, or "facts," from the medical literature with transferrable skills beyond the academy. In this article, this framework is examined through the lens of science communication ethics and writer ethos to guide individuals while navigating between automation of literature databases and human agency. Furthermore, this article explores the proper citation of research claims from different genres that are published in the "medical literature" with attention to conserving the authors' original voice. Collectively, this framework and discussion builds on prior scholarship on authorship and intellectual property in medicine.

Keywords: RHM Knowledge Integrity, Science Communication Ethics, Author Ethos, Medical Literature, Intellectual Property

## Introduction

In medicine and science, intellectual property (or "knowledge") consists of the research design, data, and insights presented by researchers in published journal articles (American Psychological Association, 2020). Such forms of intellectual property in open science, which is a social model of community and information sharing, assigns credit to the authors of the first published account of novel information (Latour & Woolgar, 1986; Katz & Linvill, 2018; Mogull, 2018a; Kapczynski, 2017; Davidoff et al., 2001). Through the lens of science communication ethics, intellectual property in scientific and medical research has been examined during the production and publishing stages through authorship and credit (Latour & Woolgar, 1986; Mogull & Katz, 2012; Katz & Linvill, 2018; Garfield, 1980). Such authorship and credit models are commonly emphasized by professional organizations such as the American Psychological Association (2020), International Committee of Medical Journal Editors (2019), and National Academy of Sciences (2009).

Despite guidance from scientific and medical associations, the actual practice of citing findings from the research literature operates in the space of writer ethos lacking institutional oversight. Notably, such practice is guided by community norms that reflect one's own research quality, subject knowledge, professional ethics, and communication intent. Thus, citing knowledge is one example of the responsibility and accountability that are underpinnings of a writer's individual agency (Lynch, 2020). Specifically, citation practice helps construct a writer's ethos through traceable and permanent record, revealing one's knowledge of and commitment to a medical discourse community. Notably, both improper attribution to the original researchers as well as misrepresentation of the rhetorical intent of published research have increasingly come into focus as a teaching or training issue, which is part of a wider gap in the communication

training of medical researchers (Derish et al., 2007). For example, in the latest review of knowledge integrity in the medical literature, nearly 15% of "facts" cited from previous studies were determined inaccurate or false, with nearly two-thirds of these quotations categorized as "major" content errors in which writers significantly mischaracterized the original research or misled their readers (Mogull, 2017). Based on the nature of the misquoted claims, scholars theorize that some writers may not be reading the full text of journal articles and therefore do not encounter the original authors' main point or conclusion (Simkin & Roychowdhury, 2005, 2007; Gavras, 2002), do not understand the information presented in the original article (Jergas & Baethge, 2015; Hartree, 1976), are intentionally misleading their audience to fit their own narrative (Greenberg, 2009; Eklund, 1995; Yankauer, 1990; Nigel Gilbert, 1977, Ingelfinger, 1976), or may misuse references to increase personal metrics of evaluation (George & Robbins, 1994; Hansen & McIntire, 1994). This relatively stable rate of 15% of falsely cited or communicated claims has permeated the medical literature for decades and has created networks of misinformation with important implications for influencing future research (Greenberg, 2009; Simkin & Roychowdhury, 2005, 2007; Ingelfinger, 1976).

Recently, a group of twenty authors of medical journal articles—apparently concerned by misquoting or mischaracterization of their own work—categorized and quantified the "factual" errors of statements by other writers who cited their work and published articles in the medical literature. Of 7,438 cited statements this group analyzed, 688 (or 9.2% of) statements were determined by the authors of the original research to be inaccurate or inappropriately cited (Pavlovic et al., 2021). As summarized in Table 1, many of the errors were interpreted by these researchers as rhetorical misuse or lack of understanding of their original research. In reflection, this group of authors emphasized that "citation inaccuracies undermine the integrity of the

scientific literature and can have serious consequences, however, good citation practices are rarely taught" (Pavlovic et al., 2021, p. 679).

Table 1. Types and frequency of incorrectly cited claims in the published medical literature as determined by the authors of the original research. Percent of errors reported from a sample of 688 cited claims communicated by authors of subsequent medical journal articles.

| | Percent of errors ($n$ = 688) |
|---|---|
| Citation of nonexistent finding | 38% |
| Inaccurately cited numerical data/results | 17% |
| Inaccurate interpretation of findings | 15% |
| Citation of quoted findings of another source (secondary citations) | 15% |
| Wrong context | 6% |
| Inaccurately cited method | 5% |
| Citation of nonexistent numerical data/results | 3% |
| Reference listed in bibliography but not cited in the text | 1% |

Note. Adapted from "How accurate are citations of frequently cited papers in biomedical literature?" by Vedrana Pavlovic, Tracey Weissgerber, Dejana Stanisavljevic, Tatjana Pekmezovic, Ognjen Milicevic, Jelena Milin Lazovic, Andja Cirkovic, Marko Savic, Nina Rajovic, Pavle Piperac, Nemanja Djuric, Petar Madzarevic, Ana Dimitrijevic, Simona Randjelovic, Emilija Nestorovic, Remi Akinyombo, Andrija Pavlovic, Ranine Ghamrawi, Vesna Garovic, and Natasa Milic, 2021, *Clinical Science*, *135*, p. 677.

Although each of these issues in Table 1 evokes concern for the field of RHM, a first step for RHM students, instructors, and practitioners is to establish a framework for systematically searching information in the medical literature that conserves the message and maintains RHM knowledge integrity. Such practice is consistent with the calls for an ethics of praxis in RHM (Melonçon, Molloy, & Scott, 2020) with implications extending beyond the medical literature to also provide patients with more accurate life-altering information, ensure the validity of public health campaigns, establish credible foundations for research funding, and promote research-informed public policy action. In this pedagogical practice article, RHM knowledge integrity is explored in the context of preparing RHM students and practitioners to be careful curators of information from the medical literature with transferrable skills beyond the academy. More

directly, the goal of this article is to provide a systematic framework for searching and citing key "facts" from the medical literature to locate "a needle in a haystack."

## Framework for Conserving RHM Knowledge Integrity while Searching and Citing the Medical Literature

The goal of this framework is to combine database automation and human agency to locate relevant research findings from the medical literature and to guide writers in the practice of conserving the rhetorical intent of the original authors. Importantly, this framework attempts to guide individuals to conserve RHM knowledge integrity and help construct the ethos of writers in the field. In this four-phase approach presented below, students and practitioners are guided through: (1) identifying an appropriate database of curated research as an empowered professional, (2) purposefully developing a strategy of keywords and filters for optimizing database searches, (3) locating and citing researchers' claims (or findings) from journal articles to conserve their voice, and (4) managing a personal library of research literature with a content management strategy that integrates time-savings and quality measures.

For instructors, this framework (covering approximately one week of instruction) may be incorporated into an existing research paper assignment or may be assigned as a short, stand-alone literature search unit focusing on theory and practice of finding and conserving the integrity of claims from the medical literature. Furthermore, this framework may be conducted individually by students or in small groups, with student work being analyzed at the end of the assignment. Alternatively, this framework may be optimally delivered as an instructor-guided series of discussions and activities as a class explores and reflects on each step in the process. As outlined in the framework and discussion below, this approach integrates practice and theory that emphasizes human agency.

**Phase 1: Identifying an Appropriate Database of Curated Research**

Identifying the database is a pivotal decision that is often taught to students from an institutional view rather than as transferrable skills to empower individuals with access to research literature beyond the institution. Specifically, undergraduate writing classes are often taught that the institutional library is the locus of information within an academic institution, which masks the actual information repository as a resource accessible outside of an intuition. To prepare students for professional careers in RHM with access to the literature outside the academy, students need exposure to and training with the disciplinary databases as resources of curated knowledge. These professionally curated databases provide a critical layer of access and quality control that are masked by broad searches of multiple academic databases merged through an institutional library interface and available outside of an institution.

Situating database selection from a practitioner's perspective reframes the database decision as a writer's agency rather than accepting the situated, institutional role as a student accepting an institutional narrative. Upon graduation, such a student accepting this lack of agency would also lack the necessary skill and, importantly, access to research the RHM literature through publicly and professionally accessible databases. Thus, we should reframe the training and selection of databases as professionally curated indexes of peer reviewed journals and teach students the names, experiences, and purposeful decisions that underly the records being searched. Furthermore, we should prepare students to access these databases outside of the institution (when possible) and position the academic library as one possible conduit for obtaining access to a specific, desired database and access to journal subscriptions.

Procedurally, the key to a successful database search of the literature is to have the research question, objective, or hypothesis drive the process. Professionally, this is first applied

to identify the academic field of research that informs the database selection as the field's accepted, curated repository of disciplinary content. For the RHM literature, two of the most relevant databases are PubMed/MEDLINE (https://pubmed.ncbi.nlm.nih.gov) that is curated and managed by the U.S. National Library of Medicine, and APA PsyInfo (typically available through intuitional libraries) that is curated and managed by the American Psychological Association. Notably, each of these professional databases has disciplinary experts that evaluate and select journals for quality metrics. Notably, these curators filter "predatory" journals from "legitimate" ones, which is as an increasing concern for reliable sources of medical research (Hansoti, Langdorf, & Murphy, 2016). Furthermore, in the case of PubMed, this is a publicly accessible database that provides practitioners and the public with free access to medical literature (Davies, 2011), which may also empower students as sources of medical research that extend postgraduation.

In contrast to disciplinary databases curated by disciplinary librarians, academic libraries and Google Scholar are broader access points that merge multiple databases with the objective of providing a single locus to initiate a search. While convenient, these resources mask the specific source of the records and undermine the effects of disciplinary curation. Practically, merged databases provide excessive, often unmanageable, output of records that cannot be carefully curated with human intention by a user. To illustrate the differences between databases, a search for "ORS in the treatment of childhood diarrhea" returned nearly 16,000 results in Google Scholar, but the same search in PubMed returned 89 results (Ramji, 2011). While both databases would return some of the same articles, the journals indexed in PubMed were "selected according [to] 'quality' criteria" (Holland, 2014), which, if removed, would require the user to reevaluate at a later stage in the search process. Thus, careful selection of an appropriate

disciplinary database can be considered the first filter of human curation applied to a literature search.

**Phase 2: Developing a Strategy of Keywords and Filters for Optimizing Database Searches**

The goal of a database search is to apply a strategic use of keywords and relevant filters to yield a manageable list of records that can be analyzed with human agency by the user in order to access all relevant content. Importantly, both the automated and human sorting of published research findings should be purposefully driven by the information needs defined at the onset of a search rather than a back-end approach of a user attempting to sort through hundreds or thousands of records. A general guideline is to target an upper limit of approximately 100 to 200 database records for human analysis, although this is an individual decision that may vary depending on external factors such as the prevalence of the published research on a topic in a particular database.

In RHM, as well as many other technical fields, the most common issue that arises is the need to narrow a list of results to produce a digestible list for the user. The primary tool applied to narrow database results is use of multiple keywords at a high level of specificity (or concreteness). Typically, these keyword filters include a combination of the condition or disease, population, test or evaluation variable, and location (Mogull, 2018b). To select and refine keyword searches, users may also access the controlled and hierarchical list of keywords used to tag articles within a database, such as the Medical Subject Headings (MeSH) for PubMed/MEDLINE (available at: https://meshb.nlm.nih.gov). Alternatively, if a keyword search yields too few results, users may reduce the number of keywords or use a broader or more inclusive term or category.

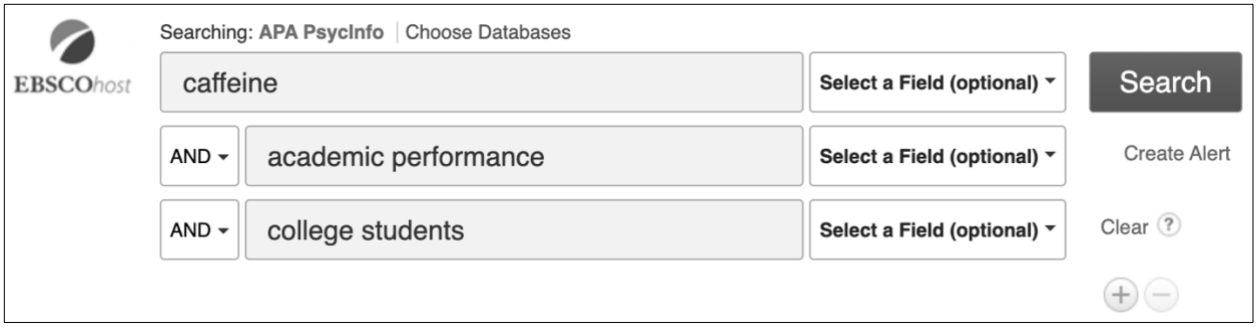
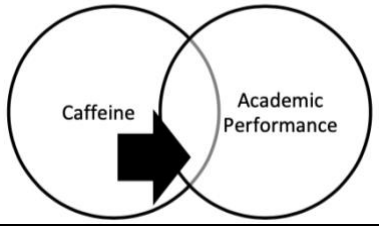
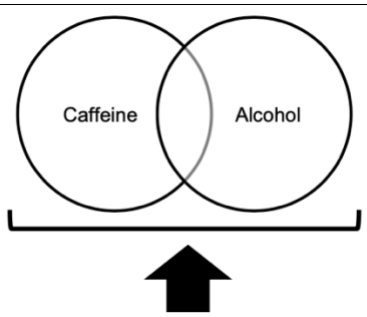
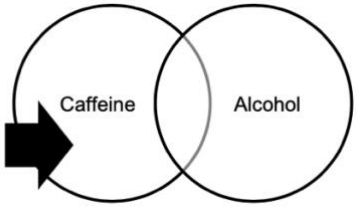
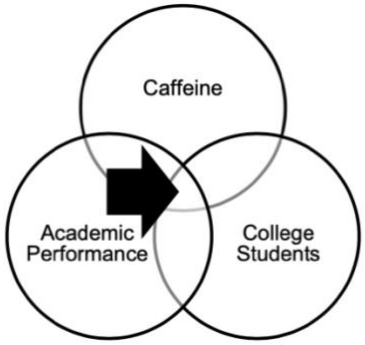



Figure 1. Advanced keyword search of APA PsyInfo using Boolean operators. In this example, multiple keywords from the research goal (examining the effects of caffeine on academic performance in college) helps narrow the database output to a manageable list of results that can be curated by humans. In this search, disciplinary research in psychology is curated in the PsyInfo database. Additional research through the lens of other disciplines, such as education or neuroscience, would supplement the search with different results.

Database searches of the research literature may be enhanced through strategic use of Boolean operators (such as AND, OR, or NOT) to selectively filter the content (Citrome, Moss, & Graf, 2009). By default, most search engines are configured to apply AND to keyword searches (see Figure 1). However, broader use of other Boolean operators (see Table 2) may be used to improve the precision of database searches. Searches can be further streamlined with a wildcard such as an asterisk (*), which returns results for any keyword prior to the truncation symbol. For example, a truncated search of "child*" would include results for child, children, and childhood.

Table 2. Boolean operators used to control and refine database results. Multiple terms and Boolean operators selected from the research question, objective, or hypothesis may be used to focus database output by content.

| **Boolean Operator** | **Use** | **Sample Keyword Search** | **Visual Representation (See indicated region)** |
|---|---|---|---|
| AND | Provides results that contain both term | Caffeine **AND** Academic Performance |  |

| | | | |
|---|---|---|---|
| OR | Provides results that contain either term | Caffeine **OR** Alcohol |  |
| NOT | Provides results from one term but excludes results that contain the other term | Caffeine **NOT** Alcohol |  |
| Applying multiple operators (e.g., AND twice) | Clarifies content sought through strategic use of multiple terms and operators | Caffeine **AND** Academic Performance **AND** College Students |  |

After receiving the initial list of records, users may consider applying additional filters to refine the search. Certain filters are more relevant and appropriate than others. For example, filters of "peer reviewed" literature would have limited impact on the results from curated databases of research journals since the source publications are curated for peer-reviewed journals. However, users may restrict the output by publication date to the most recent 1, 3, 5, or 10 years depending on the frequency of publication for a research topic to focus a search on the latest findings. Notably, some databases (such as EBSCOhost) provide a list of additional subjects or qualifiers that, as additional keywords, can be added to searches to narrow and target the results by content.

A final point to emphasize is that literature search is a continual process as research, or knowledge, is constantly being added. Depending on a writer's needs, one may need repeat database searches every 6 months to keep current with the latest medical research (Bartels, 2009). However, the frequency of repeating a literature search depends on the pace of research in a particular field and, more importantly, should be performed near the writing and publication of RHM communications. Specifically, writers are encouraged to repeat database searches of the research literature before each major revision of a manuscript and then, most importantly, before submission of the final draft for publication (Mogull, 2018b).

Since database searches are commonly repeated throughout projects, writers may consider documenting the database search parameters to efficiently repeat productive searches in the future and avoid repeating nonproductive searches. Notably, many databases allow users to create accounts and save searches for future use. Alternatively, writers may record this information in a table format (see Appendix A), which is also useful for reexamining the search strategy in reflective analysis and class discussions.

**Phase 3: Locating and Citing Claims from Journal Articles to Conserve the Authors' Voice**

To properly cite information from the medical literature, writers must read the full text of an article and, in particular, must cite the original insights or conclusion presented by the original authors of the research. Locating the proper information depends on the genre of the article, which, in the "peer-reviewed" literature, should be either primary or secondary research. *Primary research* is research authored by the original researchers of the study, which may be further subdivided into different categories to further help writers weigh the confidence of the findings. The broad categories include:

- *Basic (or theoretical) research*, which is exploratory in nature and typically provides variable levels of confidence depending on the approach and amount of evidence,
- *Observational research* (such as a case study) is designed to study one or more cases to characterize a medical condition or disease, which confers limited generalizability, and
- *Interventional research* (such as a randomized clinical trial) introduces an experimental treatment to one group for comparison to another group and confers greater confidence than observation research studies.

In these Introduction, Methods, Results, and Discussion (IMRAD)-structured journal articles, the original authors' conclusions of the study are found at the end of the article at the beginning of the Discussion or Conclusions sections. In general, these are the only statements that may be summarized and cited when presenting the main findings of an original research study. A tool for evaluating research and citing statements is provided in Appendix B.

Other genres of articles in "peer-reviewed" journals include literature review articles (or *secondary research*) that summarizes the original work of others and commentaries or opinions that express an individual's perspective on an issue. These genres should be critically evaluated as source material since these genres are rarely authored by the original researchers and, therefore, are less familiar with the original data. Furthermore, although these genres may also be "peer reviewed," the peer reviewers of secondary research do not have the original experimental data to evaluate the claims as is the case for the primary literature. Therefore, ethical practice mandates that writers must read and cite the *original, primary research* article as sources for summarizing and citing claims (International Committee of Medical Journal Editors, 2019). Such practice prevents the unethical propagation of inappropriate attribution and, most importantly, rhetorical bias that arises when writers cite secondary reports or interpretations of other journal

articles. In general, secondary research (typically labeled as “review” in the title or first page) may be used as an introduction to a topic and as guide to original research articles to access and read in full.

Not only should the statement be properly sourced and summarized, but, as illustrated in Figure 2, writers should place in-text citations immediately following the claim so that the source is clearly associated with each statement (Rivkin, 2020). If multiple articles make the same claim (or come to the same conclusion), then multiple articles may be listed following the claim (see Figure 2b). In general, writers should present the latest research findings (generally published within the past 5 to 10 years in medicine), but they should also include proper credit to older, foundational studies to ensure that credit for the original discovery or observation is assigned to the original journal article and research team (Grant, 2001; Latour & Woolgar, 1986).

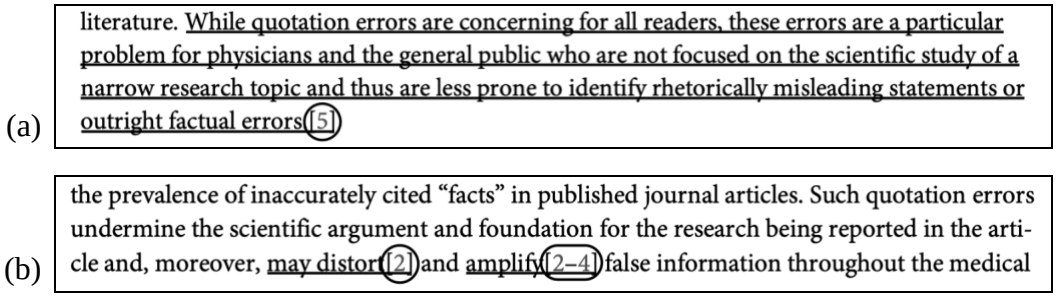



Figure 2. Placement of in-text citations to mark the cited claim. (a) Claim from a journal article summarized in a full sentence is cited at the end of the sentence. (b) Claims from multiple journal articles are cited immediately following the associated claim.

In addition to searching for the latest research, writers also must check the latest version of the journal article before citing statements and submitting their own work in order to ensure they are sourcing the latest document that does not have any corrections or revisions (Luo et al., 2013; Steen, 2011; Friedman, 1982). An increasingly visible tool is the CrossMark system that allows researchers to click on the logo from the article to check the CrossMark database for any updates (see Figure 3).

(a) 

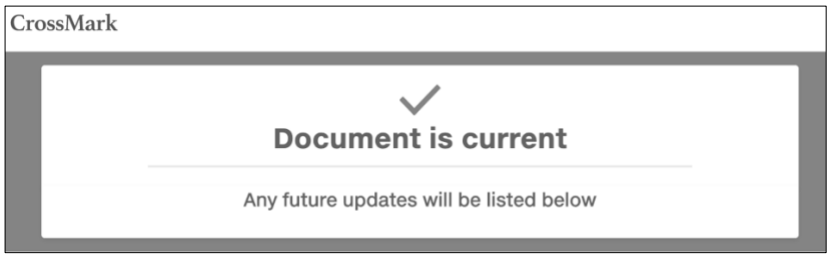
(b) 


Figure 3. Use of the CrossMark logo on medical journal articles is used to check for updates, corrections, or full retractions of information in the published literature. (a) By clicking on the CrossMark "Check for Updates" hyperlink embedded in the electronic version of journal articles, (b) the CrossMark database provides the latest summary of the content integrity of an article.

**Phase 4: Managing a Personal Library of Research Literature with a Content Management Strategy**

Several techniques are used by writers to maintain and update a personal library of relevant literature. Such personal libraries are particularly useful for writers, who as they become experts on a topic, are then able to quickly search their own database rather than repeat the searches outline above. Many writers use citation-management software, such as, RefWorks, EndNote, Mendeley, or Zotero, for managing a personal library of journal articles (Welch Medical Library, Johns Hopkins University, 2020). The major features of such software, which vary by vendor, are to provide an organized interface that writers may use to store, annotate, and retrieve articles from their desktop or online cloud drive. Furthermore, several of these systems have features that integrate into Microsoft Word so that in-text citations and lists of references may be generated automatically. Thus, such content management tools provide both time savings and quality control of citation format, which is a useful and practical skill to include when teaching literature search.

Other database tools that provide efficiency are automatic notifications of newly published content, which may be configured in many databases from prior searches or by

establishing alerts for specific articles that a user defines (Citrome, Moss, & Graf, 2009; Luzon 2007). Writers may also visit journals webpages to automatically email an electronic table of contents (eTOCs) to keep current with the research in the major journals in a field. As emphasized throughout this section, the goal is to strategically apply database tools to perform many of the presorting and time-consuming activities possible to automate in order to maximize the time and quality of attention for the necessary human agency.

## Insights and Implications for Students and Instructors

Transforming prior experience of academic literature search means that many students will retain artifacts from prior search and experience. Therefore, this information search framework was evaluated following instruction in a case study of an online, upper-division scientific writing service course situated in public research university of approximately 38,000 students. This was a purposeful sample of 16 students enrolled in Spring 2021 during the COVID pandemic (Conklin & Hayhoe, 2010; Koerber & McMichael, 2008; Goubil-Gambrell, 1992). The framework for researching and citing claims from the medical literature was evaluated through a qualitative, exploratory "teacher-research" approach for points where the framework necessitated further emphasis or clarification (Blakeslee, 2001). The primary objective of this initial analysis was to begin the reflective practice of teaching search and citation of the medical literature in the academy following interdisciplinary calls to advance instruction of medical and scientific communication (Pavlovic et al., 2021; Derish et al., 2007; Mogull, 2008).

The research of the framework employed a qualitative, mixed-methods approach using content analysis of instructor guidance followed by survey data from students completing the assignment. Initially, the framework was evaluated by a content analysis of supplemental instructional comments throughout instruction. To collect these data, all supplemental

instructional comments were recorded, transcribed, and evaluated by a content analysis for insights into student assumptions and perceptions of information search practice (Corbin & Strauss, 2014; Charmaz, 2006). Subsequently, these findings were triangulated with qualitative survey data of open-ended questions (see Appendix C) that had a response rate of 27.3% of the students completing the assignment. Research data from human subjects was collected in accordance with the standards and guidelines of the Texas State University Institutional Review Board.

Although the student sample was limited by a case study and small population of students enrolled in a single online course delivered during the COVID pandemic, and thus has implications for generalization, the purpose of this qualitative analysis was to explore complex issues within writing studies, which often employ a similar approach (Goubil-Gambrell, 1992). In the online format of this class, students independently followed the framework and, thus, a small sample of student practice was collected that provides some insight into practices of current college students throughout sciences, health sciences, and engineering fields. Collectively, these results and insights provide guidance into the literature search experiences of this generation of college students through the experiences of prevalent and seemingly similar online "Google" searches that is co-situated with literature search instruction of academic databases elsewhere in the university. To reduce the limitations of any single research method, the discussion presented below examines trends identified through both the content analysis and survey responses. Thus, this discussion may be used to provide instructors with some insights for concepts to emphasize during subsequent implementation of the information search framework.

One issue that arose is apparent disciplinary differences of database search instruction within the academy. In the humanities, many undergraduate students were trained to search the

university library database as the locus for searching academic research. In contrast, social sciences and technical disciplines show increased attention to professional and public databases, such APA PsyInfo, PubMed/MEDLINE, and the ACM Digital Library. These trends have practical implications when undergraduate students lose access to an institution's library upon graduation. Furthermore, students in some disciplines may not clearly understand the difference between implications of using disciplinary databases and the institutional host library's interface. Furthermore, using a library's database host interface at any institution lacks transference of skills beyond the academy. Promisingly, as reported in the survey, specific database recommendations from professors were internalized by undergraduate students in certain technical fields. Thus, guided literature searches by college instructors at the undergraduate level likely will influence students' future approach to locating research findings.

For RHM and other health-related research, the major difference from the research described above is that a different set of databases would be necessary to adequately search the scholarly literature in the field. For the field, which is notably interdisciplinary, purposeful selection of different databases has greater importance for a thorough search of the literature as no single database indexes content from this list of journals that publish RHM research. For example, of four popular databases examined (EBSCOhost Academic Search Complete, EBSCO Communication Source, PsyINFO, and Pub/MEDLINE), each database indexes a different corpus of journals in the field and some major journals in the field are not indexed in any of these databases (see Table 3). Therefore, keyword searches of multiple databases, including journal websites, are necessary to investigate the entire literature of RHM. In this case, University libraries provide a single portal for multiple database search and both compensate and mask the limitations of database searches in the field. Yet, University library access is limited to

those in academic institutions and students preparing for careers in non-academic settings, including policy making, would need to search multiple public databases such as PubMed/MEDLINE, Google Scholar, and journal websites for a more comprehensive search of published research literature on a specific topic.

Table 3. Journals of science, health, and medical communication/rhetoric indexed in popular databases.[1]

| | Indexed in Database | | | |
|---|---|---|---|---|
| **Journal Title** | EBSCOhost (Academic Search Complete) | EBSCO (Communication Source) | PsyINFO (APA) | PubMed/ MEDLINE |
| *Communication & Medicine* | | ✓ | | ✓ |
| *Health Communication* | ✓ | ✓ | ✓ | ✓ |
| *Journal of Business and Technical Communication (JBTC)* | | | | |
| *Journal of Health Communication* | ✓ | ✓ | ✓ | ✓ |
| *Journal of Medical Humanities* | ✓ | | ✓ | ✓ |
| *Journal of Medical Internet Research (JMIR)* | | ✓ | ✓ | ✓ |
| *Patient Education & Counseling (PEC)* | ✓ | ✓ | ✓ | ✓ |
| *Rhetoric of Health & Medicine (RHM)* | ✓ | | | |
| *Rhetoric Society Quarterly (RSQ)* | | | | |
| *Science Communication (SC)* | ✓ | ✓ | | |
| *Social Science & Medicine* | ✓ | | ✓ | ✓ |
| *Sociology of Health & Illness* | ✓ | | ✓ | ✓ |
| *Technical Communication Quarterly (TCQ)* | ✓ | ✓ | ✓ | |
| *Written Communication (WC)* | ✓ | ✓ | ✓ | |

[1] Note: List of journals indexed in databases searched on October 23, 2021. A more complete list of journals in RHM is available at: http://medicalrhetoric.com/resources/journals.

Selection and use of a database host interface is further compounded by the restrictions that students place upon the keyword selection and general avoidance of filters and library services that would improve quality or specificity of results. In contrast, undergraduate students nearly universally needed guidance on the use of filters and library services that would construct, not diminish, their ethos as they locate the most relevant research findings. Based on a content analysis of instructor comments, the practice and priorities of most searches by undergraduate students were limited to use of nouns as keywords (ranging from 1 to 3) and overall absence of a complete research question, objective, or hypothesis to narrow the search. Additionally, the students had a nearly universal lack of complex Boolean operators or nesting patterns of

keywords (including use of punctuation) to filter and focus results. Rather, the many students applied premature use of filters at the onset to thousands of records prior to human analysis, which appeared to be influenced by convenience of accessing the article rather than selecting research based on content. Notably, such a subscription issue is becoming less of an issue of convenience as full text, open-access publication models (commonplace in medicine) and relative speed of interlibrary loan programs make the immediate access of full text articles an unnecessary requirement that may have implications for knowledge integrity of citate information and the construction of personal ethos.

From the survey data, undergraduate students reported that they conducted literature searches using "phrase keywords as well as short descriptive ones like 'infection control' or 'handwashing'" and "I make sure I filter only peer-review publications, more recent years, and scientific journals." Notably, several students in the technical fields applied filters that were based on using a library interface (for example, "peer reviewed" or "scientific journals") as keywords from assignment instructions, yet the exigency of these filters were disconnected from an overall search strategy. Furthermore, many students conducted literature searches using an *information foraging* approach in which they identified a series of keywords and were guided by the database output, which led them through various, meandering paths that could lead to only a near match to the actual informational need (Greenberg, 2009; Simkin & Roychowdhury, 2005, 2007; Pirolli & Card, 1999; Bates, 1989). Therefore, undergraduate students would benefit from guided instruction determining the information needs and strategic configuration of a database search upfront to leverage database capabilities and facilitate their literature search. An important point for emphasis is the conceptual search strategy begins a writer's agency through purposeful selection of databases, keywords, and filters.

Based on the survey comments, students appeared to have sufficient prior experience with performing similar database searches and, therefore, UX with host interfaces was not identified as an issue. Rather, many students were overconfident in their skill and had distorted views of a "successful" database search. For example, one student commented that database research is "very intuitive" and another student reported, "I learned to use the databases on my own time. Each database fits my needs perfectly." Yet, these views are in contrast to conventional database search objectives for narrowing results for human review. Rather, many students believed that the higher numbers (in the thousands of tens of thousands of results) were "better" database search output and that their experience was "more successful." Yet, this high number of results is in excess of one's ability (notably time and, for undergraduate students, expertise) required for thorough analysis and selection of each result. Some unanswered questions that arise at this point are: how do students select content from this number of results, how well do those selections match the original information needs, and how much relevant information is discarded? Importantly, literature searches by current undergraduate students appear to lack the strategic use of keywords and configuration of the database fields to sort journal articles based on informative content, which would produce a highly filtered corpus of literature that may be more carefully curated and analyzed by human agency. To reframe the current view and expectations of undergraduate students while performing database searches of the research literature, the process might be positioned as a golf metaphor in which lower scores and more precise targeting of database output (through database selection, keyword strategy, and relevant filters) result in higher quality of information for citation and less overall work. Furthermore, the chances of missing relevant and important research diminishes when students

are unable to manually curate hundreds, thousands, or hundreds of thousands of database records.

Once the students obtained the database results, their final selection of research was influenced by articles immediately available through the library subscription. Many undergraduate students did not need additional guidance beyond the framework for identifying journal articles of primary research in medical or scientific fields. However, approximately 30% needed further clarification from the instructor with regards to identifying the appropriate genre and subsequent appropriateness of information for citation. Several students struggled with differentiating the secondary literature (or, "literature review articles") from the primary literature even though the word "review" was prominently located on the first page of the articles (either in the title that included the phrase "A Systematic Review" or with the genre marked as "Review" as a label on the first page). Just under 10% of the students also did not initially distinguish primary literature from commentaries, opinion articles, or the editor's introduction to a journal issue. Upon reflection, genre analysis within the medical literature needs to be further emphasized at the advanced undergraduate level.

Broadly speaking, most undergraduates were able to independently follow the framework to identify the original authors' conclusion from the primary literature (although this was dependent on proper identification of the primary literature). As summarized in the survey, one student stated,

> The discussion and results sections of the journal articles are the most appropriate sections to cite, because the recorded data and analysis of the data are found here. The abstract, introduction, and methods are not appropriate sections to cite,

because they are used to summarize the findings, introduce hypotheses, and explain how the data was collected.

However, the professional style of summary/quotation of the medical and scientific literature is notably different than the academic, lower-division practice of full article summaries. Specifically, previous undergraduate student training emphasized longer summaries that included the topic, methods, and reason of an article (often with direct quotations), whereas professional style prevalent in the medical literature includes concise, poignant restructuring of the original authors' own phrases from conclusion statements. Thus, this lesson may be expanded by tracking down additional research studies and rhetorically evaluating the style and language of quotations of a journal article that was analyzed.

Ultimately, recordkeeping and managing a personal library of medical and scientific literature were low priority skills for undergraduate students, which seems to contrast with the professional needs of students interested in researching or advocating for RHM topics as professionals. Perhaps, students may not take this extra step due to lack of exposure and experience with useful software for personal content management of the research literature. Furthermore, the demand on undergraduate students to write a wide variety of research papers on varying topics undermines the immediate value of this step as it would appear to be unnecessary and time consuming. However, the time-savings process of downloading, highlighting, notetaking, and citing research that can be stored, organized, and streamlined by citation management systems would provide RHM students with a foundation of literature that they may retain for future work on a topic. With continual emphasis on transferrable skills, citation management systems may best serve students if explored outside of institutional accounts. From an empowered writer's lens, these citation management systems provide personal permanent

access to the source material and provide automation for construction of ethos through proper in-text and reference section citation.

## Conclusions

The assignment in this case study was conducted individually to support individual research interests and many undergraduate students appeared to rely on traditional literature search practices and misinterpreted the inverse correlation between quantity and quality of results. To emphasize the salient points through the lens of human agency and developing writer's ethos, this framework might be best presented through full-class demonstration and discussion as instructors and students explore the implications of each phase of the process. Through this recommended approach, particularly relevant for interdisciplinary research such as RHM, the framework (see Appendix A and B) becomes a discussion-based, instructor-guided tutorial that is similar to the mentorship model of science and medical research training. Furthermore, rather than collecting research output, student comprehension of the principles may be better evaluated through reflection using the survey questions (see Appendix C) as a foundation for an assignment. Such an approach, possibly supplemented with this article, would emphasize agency, ethics, and strategy rather than the procedural output of a literature search assignment. In fact, many of the concepts presented formally in the framework may not resonate until reflecting and reevaluating the experience with information from the insights and implications section. As a result of using this instructor-guided approach, instructors may continue to realign the framework for different groups of students and changing outside instructional forces. Most importantly, the goal of this framework is to develop rhetorical appreciation for a process that differs from programmed prior learning.

**Scott A. Mogull** is an associate professor of technical communication at Texas State University and is the Book Review Editor for *Technical Communication Quarterly*. He has a Ph.D. in Technical Communication & Rhetoric from Texas Tech University where he researched pharmaceutical marketing. Prior to entering academia, he was a manager in the biodefense and biotechnology industries.

## References

American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.). American Psychological Association.

Bartels, Else M. (2009). How to keep up with medical literature. *Clinical Rheumatology*, *23*(2), 281–290.

Bates, Marcia J. (1989). The design of browsing and berrypicking techniques for the online search interface. *Online Review*, *13*(5), 407-424.

Blakeslee, Ann M. (2001). Bridging the workplace and the academy: Teaching professional genres through classroom-workplace collaborations. *Technical Communication Quarterly*, *10*(2), 169-192.

Charmaz, Kathy. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage Publications.

Citrome, Leslie, Moss, Stuart V., & Graf, Chris. (2009). How to search and harvest the medical literature: Let the citations come to you, and how to proceed when they do. *International Journal of Clinical Practice*, *63*(11), 1565-1570.

Conklin, James, & Hayhoe, George F. (Eds.). (2010). *Qualitative research in technical communication*. Routledge.

Corbin, Juliet, & Strauss, Anselm. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage Publications.

Davidoff, Frank, DeAngelis, Catherine D., Drazen, Jeffrey M., Hoey, John, Hojgaard, Liselotte, Horton, Richard, Kotzin, Sheldon, Nicholls, M. Gary, Nylenna, Magne, Overbeke, A. John. P.M., Sox, Harold C., & Van Der Weyden, Martin B., & Wilkes, Michael S. (2001). Sponsorship, authorship, and accountability. *The New England Journal of Medicine*, *345*(11), 825-827.

Davies, Karen S. (2011). Physicians and their use of information: A survey comparison between the United States, Canada, and the United Kingdom. *Journal of the Medical Library Association: JMLA*, *99*(1), 88-91.

Derish, Pamela A., Maa, John, Ascher, Nancy L., & Harris, Hobart W. (2007). Enhancing the mission of academic surgery by promoting scientific writing skills. *Journal of Surgical Research*, *140*(2), 177-183.

Eklund, Jan. (1995). Proper and accurate use of references. *Acta Anaesthesiologica Scandinavica*, *39*(5), 575-576.

Friedman, Barbara A. (1982). Copyright from a permissions person's point of view. *Journal of Chemical Information and Computer Sciences*, *22*(2), 70-72.

Garfield, Eugene. (1980). Has scientific communication changed in 300 years? *Current Contents*, *4*(8), 394-400.

Gavras, Haralambos. (2002). Inappropriate attribution: The “lazy author syndrome.” *American Journal of Hypertension*, *15*(9), 831.

George, Pierre M., & Robbins, Kathryn. (1994). Reference accuracy in the dermatologic literature. *Journal of the American Academy of Dermatology*, *31*(1), 61-64.

Goubil-Gambrell, Patricia. (1992). A practitioner's guide to research methods. *Technical Communication*, *39*(4), 582-591.

Grant, Maria M. (2001). Reading critically: Getting what you need from scientific papers. *Canadian Journal of Medical Laboratory Science*, *63*(1), 18-24.

Greenberg, Steven A. (2009). How citation distortions create unfounded authority: Analysis of a citation network. *BMJ*, *339*, b2680.

Hansen, Margaret E., & McIntire, Donald D. (1994). Reference citations in radiology: Accuracy and appropriateness of use in two major journals. *American Journal of Roentgenology*, *163*(3), 719-723.

Hansoti, Bhakti, Langdorf, Mark I., & Murphy, Linda S. (2016). Discriminating between legitimate and predatory open access journals: Report from the International Federation for Emergency Medicine Research Committee. *Western Journal of Emergency Medicine*, *17*(5), 497-507.

Hartree, Edward F. (1976). Ethics for authors: A case history of acrosin. *Perspectives in Biology and Medicine*, *20*(1), 82-91.

Holland, Matt. (2014). Re: What is the best web site to find citations of our research publications? https://www.researchgate.net/post/What-is-the-best-web-site-to-find-citations-of-our-research-publications/5324de53d3df3eea018b466d/citation/download

Ingelfinger F. J. (1976). Seduction by citation. *The New England Journal of Medicine*, *295*(19), 1075-6. https://doi.org/10.1056/NEJM197611042951911

International Committee of Medical Journal Editors. (2019). Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals. http://www.icmje.org/recommendations/

Jergas, Hannah, & Baethge, Christopher. (2015). Quotation accuracy in medical journal articles—a systematic review and meta-analysis. *PeerJ*, *3*, e1364.

Kapczynski, Amy. (2017). Order without intellectual property law: Open science in influenza. *Cornell L. Rev.*, *102*, 1539-1648.

Katz, Steven B., & Linvill, C. Claiborne. (2018). Lines and fields of ethical force in scientific authorship: The legitimacy and power of the office of research integrity. In Han Yu & Kathryn M. Northcut, *Scientific communication: Practices, Theories, and Pedagogies* (pp. 39-63). Routledge.

Koerber, Amy, & McMichael, Lonie. (2008). Qualitative sampling methods: A primer for technical communicators. *Journal of Business and Technical Communication*, *22*(4), 454-473.

Latour, Bruno, & Woolgar, Steve. (1986). *Laboratory life: The construction of scientific facts*. Princeton University Press.

Luo, Ma, Charles Chuan, Domingo Molina IV, Clark R. Andersen, & Panchbhavi, Vinod K. (2013). Accuracy of citation and quotation in foot and ankle surgery journals. *Foot & Ankle International*, *34*(7), 949-955.

Luzon, Maria Jose. (2007). The added value features of online scholarly journals. *Journal of Technical Writing and Communication*, *37*(1), 59-73.

Lynch, John A. (2020). The ethics of rhetoric is the rhetoric of ethics: Refusing the call to codification. *Rhetoric of Health & Medicine*, *3*(3), 249-257.

Melonçon, Lisa, Molloy, Cathryn, & Scott, J. Blake. (2020). Ethics in praxis: Situational, embodied, relational. *Rhetoric of Health & Medicine*, *3*(4), 430-436.

Mogull, Scott A. (2008). A call for new courses to train scientists as effective communicators in contemporary government and business settings. *Technical Communication*, *55*(4), 357-369.

Mogull, Scott A. (2017). Accuracy of cited "facts" in medical research articles: A review of study methodology and recalculation of quotation error rate. *PLOS ONE*, *12*(9): e0184727. https://doi.org/10.1371/journal.pone.0184727

Mogull, Scott A. (2018a). Science vs. science commercialization in neoliberalism (extreme capitalism): Examining the conflicts and ethics of information sharing in opposing social systems. In Han Yu & Kathryn M. Northcut, *Scientific communication: Practices, Theories, and Pedagogies* (pp. 64-81). Routledge.

Mogull, Scott A. (2018b). *Scientific and medical communication: A guide for effective practice*. Routledge.

Mogull, Scott A., & Katz, Steven B. (2012). *Science communication ethics vs. ethics in science* [Conference presentation]. Association of Teachers of Technical Writing, St Louis, MO.

National Academy of Sciences (2009). *On being a scientist: A guide to responsible conduct in research* (3rd ed.). The National Academies Press.

Nigel Gilbert, G. (1977). Referencing as persuasion. *Social Studies of Science*, *7*(1), 113-122.

Pavlovic, Vedrana, Weissgerber, Tracey, Stanisavljevic, Dejana, Pekmezovic, Tatjana, Milicevic, Ognjen, Lazovic, Jelena Milin, Cirkovic, Andja, Savic, Marko, Rajovic, Nina, Piperac, Pavle, Djuric, Nemanja, Madzarevic, Petar, Dimitrijevic, Ana, Randjelovic, Simona, Nestorovic, Emilija, Akinyombo, Remi, Pavlovic, Andrija, Ghamrawi, Ranine, Garovic, Vesna, & Milic, Natasa. (2021). How accurate are citations of frequently cited papers in biomedical literature? *Clinical Science*, *135*(5), 671-681.

Pirolli, Peter, & Card, Stuart. (1999). Information foraging. *Psychological Review*, *106*(4), 643-675.

Ramji, Siddarth. (2011). Effective online search for medical literature. *Indian Journal of Medical Specialities*, *2*(2), 189-191.

Rivkin, Anastasia. (2020). Manuscript referencing errors and their impact on shaping current evidence. *American Journal of Pharmaceutical Education*, *84*(7), 877-880.

Simkin, Mikhail V., & Roychowdhury, Vwani P. (2005). Stochastic modeling of citation slips. *Scientometrics*, *62*(3), 367-384.

Simkin, Mikhail V., & Roychowdhury, Vwani P. (2007). A mathematical theory of citing. *Journal of the American Society for Information Science and Technology*, *58*(11), 1661-1673.

Steen, R. Grant. (2011). Retractions in the medical literature: Who is responsible for scientific integrity? *AMWA Journal*, *26*(1), 2-7.

Welch Medical Library, Johns Hopkins University. (2020). https://browse.welch.jhmi.edu/citation-management

Yankauer Alfred. (1990). The accuracy of medical journal references. *CBE Views*, *13*(2), 38-42.