The Effectiveness of Library Instruction for Graduate/Professional Students: A Systematic Review and Meta-Analysis

Adelia Grabowsky and Liza Weisbrod

Volume 15, Number 2, 2020

URI: https://id.erudit.org/iderudit/1088816ar
DOI: https://doi.org/10.18438/eblip29657

Article abstract

Objective - This study sought to assess the effectiveness of library instruction for increasing information literacy skills and/or knowledge among graduate and professional students.

Methods - A search was conducted in Library Literature and Information Science Index (H. W. Wilson); Library, Information Science & Technology Abstracts; Medline; CINAHL; ERIC; Library and Information Science Abstracts (LISA); and ProQuest Dissertations and Theses Global. Studies were included if they were published between 2000 and 2019, in English, reported on library instruction for graduate or professional students, and objectively measured change in information literacy knowledge/skills.

Results - Sixteen studies were included in the systematic review; 12 of the 16 studies included sufficient information to be included in the meta-analysis. The overall effect of library instruction was significant [SMD = 1.03, SE=0.19, z=5.49, P<.0001, 95% CI=0.66-1.40], meaning that on average, a student scored about one standard deviation higher on an information literacy assessment after library instruction. High heterogeneity indicated a need for subgroup analysis, which showed a significant moderation of effect by discipline of students, but none by format of instruction. However, subgroup analysis must be viewed with caution due to the small number of studies in several of the subgroups.

Conclusions - This meta-analysis indicates that library instruction for graduate students is effective in increasing information literacy knowledge and/or skills. However, to strengthen the accuracy of results of future meta-analyses, there is a need for more precise descriptions of instructional sessions as well as more complete data reporting by authors of primary studies. There is also a need for the publication of more studies, particularly studies of hybrid and online instruction.
Review Article

The Effectiveness of Library Instruction for Graduate/Professional Students: A Systematic Review and Meta-Analysis

Adelia Grabowsky
Research & Instruction Librarian
Auburn University Libraries
Auburn, Alabama, United States of America
Email: abg0011@auburn.edu

Liza Weisbrod
Research & Instruction Librarian
Auburn University Libraries
Auburn, Alabama, United States of America
Email: weisbel@auburn.edu

Received: 1 Oct. 2019  Accepted: 20 Dec. 2019

© 2020 Grabowsky and Weisbrod. This is an Open Access article distributed under the terms of the Creative Commons-Attribution-Noncommercial-Share Alike License 4.0 International (http://creativecommons.org/licenses/by-nc-sa/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly attributed, not used for commercial purposes, and, if transformed, the resulting work is redistributed under the same or similar license to this one.

DOI: 10.18438/eblip29657

Abstract

Objective - This study sought to assess the effectiveness of library instruction for increasing information literacy skills and/or knowledge among graduate and professional students.

Methods - A search was conducted in Library Literature and Information Science Index (H. W. Wilson); Library, Information Science & Technology Abstracts; Medline; CINAHL; ERIC; Library and Information Science Abstracts (LISA); and ProQuest Dissertations and Theses Global. Studies were included if they were published between 2000 and 2019, in English, reported on library instruction for graduate or professional students, and objectively measured change in information literacy knowledge/skills.
Results - Sixteen studies were included in the systematic review; 12 of the 16 studies included sufficient information to be included in the meta-analysis. The overall effect of library instruction was significant [SMD = 1.03, SE=0.19, \( z=5.49, P<.0001, 95\% \text{ CI}=0.66-1.40\)], meaning that on average, a student scored about one standard deviation higher on an information literacy assessment after library instruction. High heterogeneity indicated a need for subgroup analysis, which showed a significant moderation of effect by discipline of students, but none by format of instruction. However, subgroup analysis must be viewed with caution due to the small number of studies in several of the subgroups.

Conclusions - This meta-analysis indicates that library instruction for graduate students is effective in increasing information literacy knowledge and/or skills. However, to strengthen the accuracy of results of future meta-analyses, there is a need for more precise descriptions of instructional sessions as well as more complete data reporting by authors of primary studies. There is also a need for the publication of more studies, particularly studies of hybrid and online instruction.

Introduction

Regional accrediting standards for colleges and universities emphasize the need for institutions to engage in effective assessment of desired student learning outcomes to substantiate results (Baker, 2002). One common learning outcome for university students is the ability to locate, evaluate, and manage information (i.e., to be information literate) (Markle, Brenneman, Jackson, Burrus, & Robbins, 2013). Although information literacy (IL) instruction should be interwoven throughout the curriculum, most academic librarians are invested in collaborating with subject faculty to provide library specific instruction to improve the IL skills of students (McGowan, Gonzalez, & Stanny, 2016) and are interested in assessing the value of that instruction. Library instruction to improve IL is often seen as essential only for undergraduates (Blummer, 2009). However, students in graduate/professional studies do not always have the requisite skills needed for graduate level study and research (Conway, 2011), which suggests they may also benefit from library instruction targeted specifically to graduate students. For example, O’Clair (2013) found that graduate students felt more prepared to tackle thesis research after taking a for-credit information literacy course.

Aims

This study includes both a systematic review and meta-analysis. The systematic review examines the current state of library instruction for graduate students and seeks to determine what formats of instruction are used, the content of instructional sessions, and how instruction is assessed. One issue with assessment of library instruction is that small sample sizes may limit the ability to identify actual change (Coe, 2002; Higgins, 2019). Meta-analysis is one way to combine the results of multiple studies to improve the statistical power and lessen the possibility of failing to identify a true difference (Shinogle, 2012; Thornton & Lee, 2000). Although a meta-analysis has been completed on the effectiveness of library instruction for undergraduates (Koufogiannakis & Wiebe, 2006), none was found for graduate/professional students. This study looks at the effectiveness of library instruction for graduate and/or professional students and if that effectiveness varies by discipline, format, or duration of instruction. Specific research questions include:
Systematic Review

• What formats are used to provide library instruction to graduate/professional students?
• What content is covered in instruction sessions for graduate/professional students?
• How is instruction for graduate/professional students assessed?

Meta-Analysis

• Does library instruction for graduate/professional students result in improved information literacy knowledge and/or skill?
• Does effectiveness of library instruction for graduate/professional students vary by format, duration, or discipline?

Methods

This study was conducted using the guidelines established in the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement (Moher et al., 2009). Concepts for library instruction, graduate/professional students, and study type along with synonyms and subject headings were searched on 11 March 2019 in Library Literature and Information Science Index (H. W. Wilson); Library, Information Science & Technology Abstracts; Medline; Cumulative Index of Nursing and Allied Health Literature (CINAHL); Education Resources Information Center (ERIC); Library and Information Science Abstracts (LISA); and ProQuest Dissertations and Theses Global (see Appendix A for search strategies). Academic libraries underwent significant changes in the late 1990’s with the advent of personal computers and electronic access to journal articles. Since those changes also affected library instruction, searches were limited to a date range of 2000 to 2019. Literature searches were also limited to English language, but no restrictions were placed on type of publication.

Eligibility Criteria

To be eligible for inclusion in this review, studies had to include instruction for graduate or professional students related to information literacy (IL) knowledge and/or skills. The instruction had to be provided wholly or in part by one or more librarians, and studies had to include a measure of change in IL knowledge/skills. Finally, studies had to include either one or more groups with a pre- and post-measure of IL knowledge/skills or both a treatment and control group with a post-assessment of IL knowledge/skills. Graduate students included students studying for a master’s or PhD in any subject area (other than library science), while professional students included any health science student working on a clinical doctorate, including medical, dental, pharmacy, veterinary, nursing, and audiology students. Synthesis studies, studies written in a language other than English, and studies involving medical residents or library science students were excluded. Additionally, studies were excluded if the measure of change in IL skills/knowledge was self-reported by students.

Study Selection

The number of studies examined at each stage of the review process are shown in Figure 1. Both authors independently examined each source, first at the title and abstract stage, then later at the full text stage. After each screening level, the authors compared individual decisions for congruence; conflicting decisions were resolved by discussion.

Data Extraction

Each author extracted data from half the studies to an Excel spreadsheet, and then checked data extracted by the other author for accuracy and completeness. Data collected included information about participants (level of study,
Figure 1
PRISMA flow diagram (Moher, Liberati, Tetzlaff, & Altman, 2009)
discipline, and geographic location), the intervention (description, duration, format, content taught, and content assessed), the assessment/test (timing, validity, and availability), and study statistics (sample size, mean, and standard deviation). Some studies did not include standard deviation but did provide individual scores. In those cases, standard deviations were calculated using Excel. Seven authors were emailed for additional data, and three replied with the requested information.

**Quality Assessment**

Quality of each included study was assessed using an instrument developed to critically appraise educational interventions (Morrison, Sullivan, Murray, & Jolly, 1999). The checklist includes nine questions addressing content, context, outcomes, study design, and methods. Both authors independently answered the nine questions for each study with ‘yes’, ‘no’, or ‘can’t tell’ and then met to compare results. Differences in answers to individual questions were settled by discussion and reference back to the article. The authors then voted to include or exclude the article based on preponderance of ‘yes’ answers with more weight given to questions 5 and 6. Those two questions addressed whether the study design was able to answer the posed question and whether the methods used were appropriately measuring the phenomena of interest. All articles except for one received ‘yes’ answers for both question 5 and 6 from both authors. That article received ‘no’ to both questions from both authors and was discarded due to quality concerns.

**Data Synthesis**

Analysis was carried out with R [version 3.5.0 (23 April 2018)] (R Core Team, 2018) using the [metafor](https://CRAN.R-project.org/package=metafor) package (Viechtbauer, 2010) (see Appendix B for data). Standardized mean difference (SMD) was the chosen effect size. SMD represents the difference in the pre- and post-intervention means divided by the pooled standard deviation (Borenstein, Hedges, Higgins, & Rothstein, 2009). When a study includes a small number of samples, SMD may be biased; therefore, SMD with a correction factor (Hedge’s g) was used (Borenstein et al., 2009). Hedge’s g is computed in R using the SMDH (standardized mean difference with heteroscedastic population variances in the two groups) function; SMDH requires input of the sample size, pre- and post-mean (M) scores, and pre- and post-standard deviations (SD) (see Appendix C for sample R code). When SD was not provided and could not be calculated from available data, an estimate based on the average SD of all other studies was used (Furukawa, Barbui, Cipriani, Brambilla, & Watanabe, 2006). All meta-analyses were conducted using a random-effects model, which assumes that there is not one “true” value that all studies are seeking, but instead that values may vary among studies due to differences in how the studies are carried out (Bown & Sutton, 2010). A random-effects model is recommended when there is assumed to be heterogeneity in outcome estimates (Bown & Sutton, 2010), which was the case in this meta-analysis. The I² statistic was used to quantify the heterogeneity of effect sizes. I² ranges from 0 to 100%; Borenstein et al. (2009) suggest that a small I² (close to 0) indicates that only a small part of the observed variance reflects actual differences in effect size. However, larger numbers indicate a larger proportion of the observed variance is real and suggest a need to carry out subgroup analysis or meta-regression in order to explain the heterogeneity (Borenstein et al., 2009).

**Publication Bias**

Publication bias can arise through several means, for example, authors may decide not to report non-significant findings, or journals may refuse to publish negative studies. Since meta-analysis depends on finding all studies that answer a specific research question, publication bias has the potential to distort findings (Song,
Hooper, & Loke, 2013; Thornton & Lee, 2000). This study followed the recommendations of Song et al. (2013), using a comprehensive search that did not limit results to only journal articles. In addition, publication bias was assessed with both a funnel plot and through use of Rosenthal’s fail-safe number. The fail-safe number is an estimate of the number of non-significant studies required to nullify the results of the meta-analysis; Rosenthal (1979) suggests a fail-safe number greater than $5n + 10$ (where $n$ is the number of studies) is sufficient to consider publication bias inconsequential (see Appendix C for sample code to calculate Rosenthal’s fail-safe number in R).

**Results**

**Description of Studies**

The final 16 studies included one dissertation and 15 journal articles, with publication years ranging from 2004 to 2018 (see Appendix D for list of studies and Table 1 for characteristics of studies). The majority of studies took place in the United States ($n=12$), but there was one study each from Canada, the United Kingdom, Australia, and Tanzania. A variety of disciplines were represented (see Figure 2), with the largest number of studies including medical students ($n=4$) followed closely by students in education ($n=3$). Overall, health science students were included more often, with 10 of 16 studies involving students from some area of health sciences. The 16 studies included 12 studies that were pre- and post-assessments of one or more groups (repeated measures); the remaining four were post-assessment of a treatment and control group (independent groups). Sample size of the repeated measures (RM) studies ranged from 10 to 61 students, while sample size of the independent group (IG) studies ranged from 37 to 300 students. The most common format of instruction was face-to-face (F2F). Six studies included only F2F instruction, five included only hybrid instruction (a combination of face-to-face and some sort of online instruction), and two studies examined only online instruction. An additional two studies compared F2F to online, and one study compared all three formats--F2F, hybrid, and online. Duration of instruction was not reported for every study. Durations that were reported varied widely; for example, for library instruction provided within a subject class, time of instruction ranged from one 70-minute session to two 3-hour sessions (see Table 2).

**Content of instruction**

While not every study included a detailed description of instructional content, certain themes emerged in the studies (see Table 1). All classes taught database searching strategies ($n=16$). The classes for health sciences students ($n=10$) provided instruction on biomedical databases (PubMed, International Pharmaceutical Abstracts, MEDLINE, CINAHL, and others) while classes for non-health science students ($n=6$) taught a variety of resources including chemistry and education databases and sources of data from United States government agencies. Search strategies taught included Boolean logic ($n=6$), limiters ($n=5$), and MeSH (Medical Subject Headings) vocabulary ($n=5$). Other topics included critical appraisal skills ($n=6$), citation styles and citation managers ($n=4$), ethical use of information (including plagiarism) ($n=3$), and library-specific resources and services ($n=4$).

**Assessment**

Six of the 16 studies employed a validated assessment tool (see Table 3). Of those six, two used an instrument based on the Fresno test (Ramos, Schafer, & Tracz, 2003), one used the RRSA (Research Readiness Self-Assessment) (Ivanitskaya, Laus, & Casey, 2004), and two studies used rubrics validated in-house. The remaining study used backwards design and the Information Literacy Competency Standards for Higher Education (Association of College and Research Libraries [ACRL], 2000) to develop a validated assessment tool. Nine studies
<table>
<thead>
<tr>
<th>Shortened Citation</th>
<th>Participants Level Discipline Location</th>
<th>Design Repeated measures=RM Independent groups=IG</th>
<th>Format F2F (Face to Face) Online Hybrid</th>
<th>Description of intervention Duration of intervention</th>
<th>Content Taught Content Assessed</th>
</tr>
</thead>
</table>
| Aronoff, 2017 Teaching evidence-based practice... | Students from 8 health profession programs (medical, dental, pharmacy, occupational therapy (OT), physical therapy (PT), social work, speech language pathology, dietetics) USA | RM Pre/Post-assessment 1 Tx group (n=39) | Hybrid | • 2 online evidence based practice (EBP) learning modules hosted on the learning management system.  
• Participation in a facilitated in-person interprofessional small group learning experience.  
• Each module 1 hour long. | Taught: Module 1: EBP principles, critical appraisal strategies. Module 2: PubMed instruction, Medical Subject Headings (MeSH) terms.  
Assessed: Module 1: Knowledge of EBP components; development of patient/population, intervention, comparison, and outcome (PICO) questions; study designs; critical appraisal strategies. Module 2: PubMed searching strategies; using MeSH terms; limiting with PubMed filters. Clinical scenario: creation of a PICO question, utilization of information resources, study design, search characteristics, and critical appraisal. |
<p>| Beile, 2004 Does the medium matter? | Master’s, Doctoral, and certificate-seeking education | RM Pre/Post-test 3 Tx groups | Study compared 3 modes of delivery. | Taught: F2F: demonstration of relevant library databases followed by an activity to allow the students to apply the lesson. |</p>
<table>
<thead>
<tr>
<th>Shortened Citation</th>
<th>Participants</th>
<th>Level</th>
<th>Discipline</th>
<th>Location</th>
<th>Design</th>
<th>Format</th>
<th>Description of intervention</th>
<th>Duration of intervention</th>
<th>Content Taught</th>
<th>Content Assessed</th>
</tr>
</thead>
</table>
| students            | Group 1 (F2F, on-campus) (n=16) | Repeated measures=RM | Independent groups=IG | F2F | • An on-campus class with face-to-face library instruction.  
• 70-minute demo followed by application activity. | | | |
| USA                | Group 2 (web tutorial, on-campus) (n=19) | RM | Pharmacy students 1st year | Pre/Post-test 1 Tx group (n=61) | Hybrid | • An on-campus class with Web-based library tutorial consisting of 4 interactive modules.  
• Participants spent an average of 80 minutes on modules. | | | |
| Chiarella, 2014 Information literacy skills retention… | Group 3 (web tutorial, web-based class) (n=14) | Pre/Post-test | Pharmacy students 1st year | USA | Online | • A web-based class with a web-based library tutorial consisting of 4 interactive modules.  
• Participants spent an average of 80 minutes on modules. | | | |
|                    | RM | RM | F2F | • Librarian presented library skills material 4 times during the fall semester of P1 year.  
• No indication of length of session. | | Taught: Basic database search strategies; Google searching versus biomedical databases; PubMed, EMBASE, and MEDLINE; EndNote. | | Assessed: MeSH subject heading searches, Boolean operators, and limits. |
<table>
<thead>
<tr>
<th>Shortened Citation</th>
<th>Participants Level Discipline Location</th>
<th>Design</th>
<th>Format</th>
<th>Description of intervention</th>
<th>Duration of intervention</th>
<th>Content Taught</th>
<th>Content Assessed</th>
</tr>
</thead>
</table>
| Dorsch, 2004       | Medical students 3rd year USA          | RM     | F2F    | Taught: Librarians: to define evidence based medicine (EBM), formulate clinical questions based on a standardized case scenario; identify and review EBM search strategies and resources. | 8 1-hour weekly seminars.  
• Weeks 1-2 taught by librarian.  
• Weeks 3-5 taught by medical school faculty.  
• Weeks 6-8 practice sessions. | Taught: | Assessed: Formulating a clinical question, using effective strategies to identify the best clinical literature to answer the question, analyzing the relevance and validity of the retrieved article. |
| Emmett, 2007       | 1st, 2nd year PhD Chemistry students USA | RM     | F2F    | Taught: Major resources in the chemical and biomedical literature, research strategies, bibliographic management, ethical use of information. | 1-hour credit course taught by librarian.  
• 75 minutes per week for one semester.  
| Grant, 2006        | Master’s, PhD students (Nursing, OT, PT) United Kingdom | RM | Hybrid (Online tutorial within EBP module) | Taught: Tutorial: the rationale for a literature search; how a database works; seven search steps covering clarifying a search question, breaking down the question, MeSH, free text | During a 12-week EBP module, 2 sessions (3 hours each) were allocated to information skills development.  
• An online tutorial was used in-class for both sessions, | Taught: | |
<p>| | | | | | | | |
|                    |                                        |        |        |                             |                          |                |                 |</p>
<table>
<thead>
<tr>
<th>Shortened Citation</th>
<th>Participants Level Discipline Location</th>
<th>Design</th>
<th>Format</th>
<th>Description of intervention Duration of intervention</th>
<th>Content Taught Content Assessed</th>
</tr>
</thead>
</table>
| Ilic, 2012 Teaching evidence-based medicine... | Medical students 3rd year Australia | IG | Hybrid | - EBM literature searching skills workshop (intervention group attended workshop, control group did not).  
- Workshop consisted of formal presentation by | Taught: How to construct an answerable question from the clinical environment, major sources of medical information, how to effectively and efficiently search the medical literature to |
| (n=10, extended) | | | | and students were asked to complete between-session exercises using the tutorial.  
- No indication of length of tutorial. | Assessed: Short-term, a literature search; longer term, systematic literature search on a topic of choice, describing the literature search process and providing search strategies, then selecting and critically appraising two papers. Both assessed by skills checklist such as Boolean operators, use of MeSH/indexing terms, application of limits, and whether a manageable and relevant number references were retrieved. |

**Design**
- Repeated measures=RM  
- Independent groups=IG

**Format**
- F2F (Face to Face)  
- Online  
- Hybrid

**Duration of intervention**
- IG 1 Tx group (n= 60)  
- IG 1 control group (n=37) block
<table>
<thead>
<tr>
<th>Shortened Citation</th>
<th>Participants Level Discipline Location</th>
<th>Design</th>
<th>Format</th>
<th>Description of intervention</th>
<th>Duration of intervention</th>
<th>Content Taught</th>
<th>Content Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ivanitskaya, 2008</td>
<td>Master of Science in Administration Students USA</td>
<td>RM Pre/Post-test 1 Tx group (n=14)</td>
<td>F2F</td>
<td>Library instruction during a class session at the beginning of the course. Class sessions were from 5:30 to 10 pm but amount of time given to library instruction was not specified.</td>
<td>• Library instruction during a class session at the beginning of the course. • Class sessions were from 5:30 to 10 pm but amount of time given to library instruction was not specified.</td>
<td>Taught: Search strategies (keywords, subject headings, and Boolean operators), how to find journal articles, identifying and searching for scholarly journals, searching for articles using the appropriate journal database for the topic, refining the search, evaluating the article, and downloading or ordering the full-text of the article.</td>
<td>Assessed: Ability to find information, ability to evaluate information, and understanding of plagiarism.</td>
</tr>
<tr>
<td>Shortened Citation</td>
<td>Participants Level Discipline Location</td>
<td>Design</td>
<td>Format</td>
<td>Description of intervention Duration of intervention</td>
<td>Content Taught</td>
<td>Content Assessed</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------------------</td>
<td>--------</td>
<td>--------</td>
<td>------------------------------------------------------</td>
<td>---------------</td>
<td>-----------------</td>
<td></td>
</tr>
</tbody>
</table>
| Lapidus, 2012 Combined use of online tutorials... | Pharmacy students 2nd year USA | IG | Hybrid | • Control group (2008) received library instruction using lecture and demo.  
• Intervention group (2010) used blended learning with online tutorials, brief demo, in-class hands-on exercises, group discussion.  
• 5 to 6 class sessions taught by librarians during a fall semester.  
Assessed: Answering drug information questions using tertiary print and electronic resources; searching Medline and IPA. | |
| Lechner, 2005 Graduate student research instruction... | Master of Science in Occupational Therapy and Master of Physical Therapy USA | RM | Online | • Online tutorial that provided live results in response to students’ actions.  
• Each class (OT and PT) was randomized into 2 groups; one group went to another room to complete the online tutorial while the remaining | Taught: Searching CINAHL database including controlled vocabulary, functions of various indexes, using limits to filter and focus results.  
Assessed: Basic information literacy (e.g., definition of peer-reviewed), basic CINAHL | |
<table>
<thead>
<tr>
<th>Shortened Citation</th>
<th>Participants Level Discipline Location</th>
<th>Design</th>
<th>Format</th>
<th>Duration of intervention</th>
<th>Description of intervention</th>
<th>Content Taught</th>
<th>Content Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maranda, 2016 Evaluation of the long-term impact…</td>
<td>Medical students 1st year Canada</td>
<td>RM Pre/Post-test 1 Tx group (n=100, pre) (n=59, post)</td>
<td>Hybrid</td>
<td>• Library instruction in year 1 consisted of 3 online modules and 3 in-person sessions. &lt;br&gt;• No indication of length of sessions or online modules.</td>
<td>students attended a lecture covering the same material. Students in the lecture group could choose to watch only or could follow along on computers. &lt;br&gt;• No information about length of class.</td>
<td>characteristics (e.g., target audience), basic CINAHL skills (e.g., combining searches), advanced CINAHL skills (e.g., interpreting hierarchy of subject headings), advanced CINAHL characteristics (e.g., using account to store results).</td>
<td>Taught: E-books, POC (point of care) tools, MEDLINE/PubMed searching, drug information resources. &lt;br&gt;Assessed: Only 2 questions (knowledge of Boolean logic, choice of resource for clinical scenario) were consistent across the pre- and post-tests, and the survey.</td>
</tr>
<tr>
<td>Otto, 2012 Assessing and improving data literacy…</td>
<td>Master’s students Urban Regional Planning College of Business and Public Administration USA</td>
<td>RM Pre/Post-text 1 Tx group (n=13)</td>
<td>F2F</td>
<td>• 2 sessions, each a combination of lecture and hands-on teaching. &lt;br&gt;• Each session was 90 minutes, first session was about 1/3 through semester, second was about 2/3 through semester.</td>
<td></td>
<td>Taught: Selecting and using appropriate data sources, retrieving needed data. 1st session focused on demographic and population data, 2nd session on economic data.</td>
<td>Assessed: Knowledge of trusted government sources, conceptual</td>
</tr>
<tr>
<td>Shortened Citation</td>
<td>Participants</td>
<td>Design</td>
<td>Format</td>
<td>Description of intervention</td>
<td>Content Taught</td>
<td>Content Assessed</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------</td>
<td>--------</td>
<td>--------</td>
<td>-----------------------------</td>
<td>----------------</td>
<td>-----------------</td>
<td></td>
</tr>
<tr>
<td>Schilling, 2006</td>
<td>Medical students 3rd year USA</td>
<td>IG</td>
<td>Online</td>
<td>• During weeks 1 &amp; 2 of a 6-week clerkship, students used 2 course-integrated, web-based learning modules designed by health science librarians. • Modules required 40 to 60 minutes to complete.</td>
<td>Taught: Basic MEDLINE searching; using MeSH, Boolean operators; finding randomized controlled trials, meta-analyses, &amp; gold standard literature; searching the Cochrane database; information found in different types of research.</td>
<td>Assessed: Ability to formulate a clinical question, develop an effective search strategy, ID and use correct MeSH terms, use Boolean operators, use appropriate limits, restrict search results to randomized controlled trials or meta-analyses.</td>
<td></td>
</tr>
</tbody>
</table>

An interactive web-based curriculum...
<table>
<thead>
<tr>
<th>Shortened Citation</th>
<th>Participants</th>
<th>Design</th>
<th>Format</th>
<th>Description of intervention</th>
<th>Content Taught</th>
<th>Content Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schweikhard, 2018</td>
<td>Master of Occupational Therapy &amp; Doctor of Physical Therapy students USA</td>
<td>IG</td>
<td>1 Tx group (n=90) 1 control group (n=90) participants randomly selected from a larger pool</td>
<td>Online</td>
<td>9 online library instructional tutorials created with Guide on the Side and embedded in the online course platform.</td>
<td>Taught: Overview of library; using appropriate databases; using MeSH; searching for different types of evidence (randomized controlled trials, systematic reviews and meta-analyses, cohort &amp; case-control studies, diagnostic tests, qualitative research, practice guidelines).</td>
</tr>
<tr>
<td>Shaffer, 2011</td>
<td>Curriculum and Instruction Dept., College of Education USA</td>
<td>RM</td>
<td>Pre/Post-test 2 Tx groups</td>
<td>Online</td>
<td>Online tutorial consisted of 8 mini-tutorials.</td>
<td>Taught: Sources for quality education research; scholarly and primary research; choosing search terms; searching effectively and efficiently; finding full-text, APA citation; other database features.</td>
</tr>
<tr>
<td>Shortened Citation</td>
<td>Participants Level</td>
<td>Design</td>
<td>Format</td>
<td>Description of intervention</td>
<td>Content Taught</td>
<td>Content Assessed</td>
</tr>
<tr>
<td>--------------------</td>
<td>-------------------</td>
<td>--------</td>
<td>--------</td>
<td>-----------------------------</td>
<td>----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td></td>
<td>Disciplines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Wema, 2006         | Master of Education students, Tanzania | RM | Pre/Post-test 1 TTx group (n=12) | F2F | A combination of lectures and hands-on activities.  
- 7-day course, each day began at 8 am and lasted to approximately 5 pm. | Taught: Formulating a question; defining information needs; organizing ideas for information need; categories and structure of information sources; developing search strategies, how to modify search; capturing and synthesizing information from sources; evaluating sources; presenting information; referencing and citing; ethical and legal issues in using information. | |

F2F group (n=30)  
F2F  
of students in F2F and the other half in a different room for the online tutorial.  
- Both F2F and online tutorial took place during a 3-hour class; students had instruction either F2F or online then used remaining time for independent research.  
- Online session averaged less than 2 hours, F2F instruction averaged 2 hours.  

primary research; choosing search terms; searching effectively and efficiently; finding full-text, APA citation; other database features.
<table>
<thead>
<tr>
<th>Shortened Citation</th>
<th>Participants Level</th>
<th>Discipline</th>
<th>Location</th>
<th>Design</th>
<th>Format</th>
<th>Duration of intervention</th>
<th>Content Taught</th>
<th>Content Assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Repeated measures=RM</td>
<td>F2F (Face to Face)</td>
<td></td>
<td>Assessed: Defining a problem or research topic; information sources; internet sources; internet search; library and database searching; evaluating information and sources; referencing; synthesizing information; presenting information.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Independent groups=IG</td>
<td>Hybrid</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Table 2
Duration of Instruction

<table>
<thead>
<tr>
<th>Not mentioned (n=4)</th>
<th>Online tutorials (n=3)</th>
<th>Stand-alone classes (n=2)</th>
<th>Sessions within subject classes (n=7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No mention of duration (n=2)</td>
<td>2 modules, 50 minutes total</td>
<td>75 min/week for one semester (n=1) [for credit class]</td>
<td>1 session @ 70 min (n=1)</td>
</tr>
<tr>
<td>Mentioned # of sessions but not length of sessions (n=2)</td>
<td>2 modules, 120 minutes total</td>
<td>Each day for 1 week (n=1) [seminar]</td>
<td>1 session @ 120 minutes (n=2)</td>
</tr>
<tr>
<td></td>
<td>4 modules, 80 minutes total</td>
<td></td>
<td>1 session @ 180 minutes (n=2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 sessions @ 90 minutes each (n=1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2 sessions @ 180 minutes each (n=1)</td>
</tr>
</tbody>
</table>

Figure 2
Number of studies by discipline.
provided the full questionnaire or assessment in the article.

Seven of the 16 studies referenced the Information Literacy Competency Standards for Higher Education (Association of College and Research Libraries, 2000) with one of the seven also referencing the Framework for Information Literacy for Higher Education (Association of College and Research Libraries, 2016). Of the nine remaining studies that did not reference ACRL standards, three referenced other standards; for example, the proposed Core Competencies for Data Information Literacy (Carlson, Fosmire, Miller, & Nelson, 2011).

Most studies (n=9) used objective tests such as multiple choice and true/false questions to measure recall of knowledge, while five studies measured the application of knowledge by evaluating search strategies or scenario responses. Two studies measured the application of knowledge through short answer and multiple-choice questions that required hands-on use of databases.

Meta-Analysis

Meta-analysis often involves examination of experimental studies involving independent groups (IG), for example treatment and control groups; however, meta-analysis is also possible with repeated measures designs (RM). RM studies involve one or more groups; individuals within the groups are assessed both before and after an intervention. These two types of studies differ in the type of research question involved, with IGs interested in group differences while RMs explore change at the individual level (Morris & DeShon, 2002). Morris and DeShon (2002) point out that combining IG and RM studies may be done, but only if effect sizes are transformed to account for differences in how standard deviations are calculated. The IG studies found in this systematic review did not include the information required to transform the effect sizes to equivalent RM effect sizes as recommended (Morris and DeShon, 2002). In addition, the small number of IG studies was considered insufficient to complete a separate meta-analysis, therefore only the RM studies (pre- and post-assessment of one or more groups) were included in the meta-analyses. When an RM study included multiple groups, for example, a comparison of online versus face-to-face instruction, each group was considered separately in the meta-analysis. Therefore, for the 12 RM studies there were 16 associated effect sizes. Nine of the RM effect sizes involved face-to-face instruction (F2F) by a librarian, three were online modules only, and four were hybrid sessions, involving F2F instruction supplemented with online modules.

Effectiveness of library instruction for graduate students

A meta-analysis run on all RM groups (16 effect sizes from 12 studies) produced an overall standardized mean difference (SMD) of 1.03 [SE=0.19, z=5.49, P<.0001, 95% CI=0.66-1.40], (see Figure 3), which is considered a large effect size (Cohen, 1988). Another way to state the result is that graduate students scored slightly more than one standard deviation higher on a measure of IL skills after receiving library instruction. The I² statistic for the meta-analysis was 81.47%, indicating a large amount of heterogeneity (Higgins, Thompson, Deeks, & Altman, 2003) and a need for subgroup analysis. One possibility is that results could have been influenced by the estimation of standard deviation (SD) for two studies (3 associated effect sizes) (Appendix B); however, a test of estimation as a moderator revealed no significant difference between studies with estimated variables and those without (QM (df=1) =0.24, P=.63).

Effectiveness by format, duration, or discipline

Sub-group analysis was completed in an attempt to explain the large amount of heterogeneity in the overall meta-analysis. Format of instruction
| Shortened Citation | a) Questionnaire/assessment notes  
b) Timing  
c) Tested recall of knowledge or application of knowledge? | Questionnaire/assessment  
a) Validity/reliability addressed?  
b) Used existing or created for this study?  
c) Author(s) referenced the ACRL Standards or Framework?  
d) Full questionnaire/assessment available? | Findings | Additional assessments? |
|--------------------|-------------------------------------------------|------------------------------------------------|----------|----------------------|
| Aronoff, 2017  
Teaching evidence-based practice... | a) AFT (Adapted Fresno Test).  
b) Given before students had access to the online modules then again after they completed the online modules, then a 3rd time after they participated in the small group learning experience.  
c) Application. | a) Yes.  
b) Used existing Adapted Fresno Test.  
c) No.  
d) No. | Scores on the AFT increased significantly post-modules, but decreased post-small group experience. | Students took a quiz after each of the 2 modules.  
Students completed an anonymous evaluation of the modules. |
| Beile, 2004  
Does the medium matter? | a) 20 multiple choice questions.  
b) Pre-test given immediately before the instruction session began.  
Post-test given ~ 6 weeks after the instruction session.  
c) Recall. | a) No.  
b) Created for this study, written by faculty who teach library instruction sessions.  
c) Yes, assessment based on ACRL IL Standards.  
d) No | Significant increase in post-instruction scores; no significant difference in scores by format of instruction. | Self-reported perceptions of efficacy. |
| Chiarella, 2014  
Information literacy skills retention... | a) 7-item multiple choice quiz.  
b) Pre-test administered before the 1st library instruction of P1 fall semester, post-test administered at end of P1 spring semester.  
c) Recall. | a) No.  
b) Created for this study.  
c) No.  
d) Yes. | No significant difference between pre- and post- scores. | No |
<table>
<thead>
<tr>
<th>Shortened Citation</th>
<th>a) Questionnaire/assessment notes</th>
<th>b) Timing</th>
<th>c) Tested recall of knowledge or application of knowledge?</th>
<th>Questionnaire/assessment a) Validity/reliability addressed?</th>
<th>b) Used existing or created for this study?</th>
<th>c) Author(s) referenced the ACRL Standards or Framework?</th>
<th>d) Full questionnaire/assessment available?</th>
<th>Findings</th>
<th>Additional assessments?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dorsch, 2004 Impact of an evidence-based...</td>
<td>a) Students given simulated case scenarios, which were different for pre- and post-tests. Scenarios evaluated by both a librarian and a faculty member using a competency-based instrument with 15 items, each scored from 1 to 7. b) Assessment given at beginning and end of seminar series. c) Application.</td>
<td>a) No.</td>
<td>b) Created for this study.</td>
<td>c) No (referenced Medical School Objectives).</td>
<td>d) Yes.</td>
<td>Statistically significant improvement occurred in creating a PICO question; using MeSH, Boolean, and limits; assessing articles.</td>
<td></td>
<td>Pre- and post-survey to assess students’ self-perception of change in EBM skills.</td>
<td></td>
</tr>
<tr>
<td>Emmett, 2007 Assessing information literacy skills...</td>
<td>a) 29 multiple choice/short answer questions. b) Pre-test given at beginning of semester, post-test given at end of semester. c) Recall and application.</td>
<td>a) Yes.</td>
<td>b) Created for this study.</td>
<td>c) Yes, assessment created using ACRL IL standards and backward design.</td>
<td>d) Yes.</td>
<td>57% increase in post-test scores, no statistical analysis provided.</td>
<td></td>
<td>Assessment related to class itself, including in-class exercises, final project, and final exam.</td>
<td></td>
</tr>
<tr>
<td>Grant, 2006 Developing and evaluating an interactive...</td>
<td>a) Assessment of searches. b) Assessed search done at the beginning of the first session (pre), search done at the end of the second session (post), and search done at the end of the 12 week class (extended). c) Application.</td>
<td>a) No.</td>
<td>b) Used existing assessment tool modified from Rosenberg et al. (1998).</td>
<td>c) No.</td>
<td>d) Yes.</td>
<td>Statistically significant difference between pre- and post-scores; and between post- and extended scores.</td>
<td></td>
<td>Subjective evaluation of students’ perceptions of learning.</td>
<td></td>
</tr>
<tr>
<td>Shortened Citation</td>
<td>Questionnaire/assessment notes</td>
<td>Questionnaire/assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------</td>
<td>--------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ilic, 2012 Teaching evidence-based medicine…</td>
<td>a) Fresno test. b) Post-test assessment done at 1 week post-implementation of intervention. c) Application.</td>
<td>a) Yes. b) Used existing Fresno test. c) No. d) No.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Findings**

- Ilic, 2012: No statistically significant difference in scores between the treatment and control group.
- Ivanitskaya, 2008: Statistically significant difference in pre- and post-test scores. RRSA also includes subjective measures that ask for students’ perceptions of research skills and previous library/research experience.
- Lapidus, 2012: No difference in students’ scores when comparing hybrid instruction to traditional lecture-based instruction. Additional homework assignment covering tertiary resources; course evaluations with students’ perceptions of course design and teaching methods.

**Additional assessments?**

- Clinical Effectiveness and Evidence-based Practice Questionnaire (EBPQ) used to assess students’ self-perceived competency in EBM literature searching.
<table>
<thead>
<tr>
<th>Shortened Citation</th>
<th>Questionnaire/assessment notes</th>
<th>Timing</th>
<th>Tested recall of knowledge or application of knowledge?</th>
<th>Validity/reliability addressed?</th>
<th>Used existing or created for this study?</th>
<th>Author(s) referenced the ACRL Standards or Framework?</th>
<th>Full questionnaire/assessment available?</th>
<th>Findings</th>
<th>Additional assessments?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lechner, 2005</td>
<td>a) 20 multiple choice questions delivered using WebCT. b) Pre-test given immediately before instruction began. Post-test given at a later unspecified date. c) Recall.</td>
<td>a) No. b) No indication of origin of questionnaire. c) No. d) No.</td>
<td>Average scores increased after instruction, no statistical analysis provided.</td>
<td>2 additional questions asked about students’ prior use of CINAHL.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gradate student research instruction…</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maranda, 2016 Evaluation of the long-term impact…</td>
<td>a) 5-item multiple choice test. b) Pre-test completed online in the first few days of medical school. Post-test administered at end of first year. Questions asked again at the end of fourth year. c) Recall.</td>
<td>a) Mentioned validity/reliability of assessments considered but not used, however validity/reliability of their questionnaire was not addressed. b) Created for this study, piloted with 4 medical students and 5 librarians and changes made based on feedback. c) No. d) Yes, in supplementary material.</td>
<td>Statistically significant increase in scores between pre-test and post-test; increase in scores at end of 4th year but no statistical analysis provided.</td>
<td>Post-program survey of attitudes and behaviors, and confidence in EBM tasks. Results of pre- and post-tests compared to post-program survey.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shortened Citation</td>
<td>Questionnaire/assessment notes</td>
<td>Timing</td>
<td>Tested recall of knowledge or application of knowledge?</td>
<td>Findings</td>
<td>Additional assessments?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------------</td>
<td>--------</td>
<td>--------------------------------------------------------</td>
<td>----------</td>
<td>------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Otto, 2012</td>
<td>a) 12 multiple choice/matching questions; identical pre/post-tests. b) Pre-test given before 1st session. Post-test given at end of quarter. c) Recall.</td>
<td>a) No. Feedback on the questionnaire was solicited from library colleagues; the course instructor vetted the final questionnaire. There was no trial run before use. b) Created for this study. c) No (referenced Core Competencies for Data Information Literacy). d) Yes, in Appendix.</td>
<td>Average scores increased from pre-test to post-test, no statistical analysis provided.</td>
<td>Examination of student assignments.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schilling, 2006</td>
<td>a) Analysis of students’ MEDLINE search strategies. b) Final (6th) week of rotation. c) Application.</td>
<td>a) Yes. b) Evaluation criteria developed in previous research. Interrater reliability assessed on evaluations of search strategy. c) Yes (ACRL IL standards). d) No.</td>
<td>Scores for treatment group were significantly greater than the control group.</td>
<td>Pre- and post-clerkship survey (self-report); post-clerkship NNT (Number Needed to Treat) test (self-report); analysis of articles identified as best evidence.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schweikhard, 2018</td>
<td>a) Final course papers scored by 2 independent reviewers using a rubric. b) End of course. c) Application.</td>
<td>a) Yes, two reviewers for each paper. Reviewers practiced scoring papers not selected for the assessment sample to support interrater reliability. b) Scoring rubric created for this study.</td>
<td>Statistically significant increase in post-tutorial students’ use of search terms, MeSH headings, limits, and level of evidence of cited studies. There was no increase in post-tutorial</td>
<td>No.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shortened Citation</td>
<td>Questionnaire/assessment notes</td>
<td>Timing</td>
<td>Tested recall of knowledge or application of knowledge?</td>
<td>Findings</td>
<td>Additional assessments?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------------------</td>
<td>--------</td>
<td>------------------------------------------------------</td>
<td>----------</td>
<td>------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shaffer, 2011</td>
<td>a) 20 multiple choice questions</td>
<td>b) Pre-test given immediately before instruction. Post-test varied according to instructors’ assignment schedule, given after the first iteration of the “works cited list” was due. c) Recall.</td>
<td>a) No. b) Some questions adapted from a validated test (Beile Test of Information Literacy in Education); questions addressed learning outcomes. c) Yes (ACRL IL Standards). d) Yes.</td>
<td>Statistically significant improvement in scores after instruction for both groups (F2F &amp; online); no significant difference in scores between the F2F and online students.</td>
<td>Citation analysis; five questions to determine students’ general level of confidence in key library research skills; students using online tutorial also completed quizzes after each module and a survey asking about tutorial design and comfort with technology used.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>a) Questionnaire/assessment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>notes</td>
<td>b) Timing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c) Tested recall of knowledge or application of knowledge?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shortened Citation</td>
<td>a) Questionnaire/assessment notes</td>
<td>b) Timing</td>
<td>c) Tested recall of knowledge or application of knowledge?</td>
<td>Questionnaire/assessment</td>
<td>a) Validity/reliability addressed?</td>
<td>b) Used existing or created for this study?</td>
<td>c) Author(s) referenced the ACRL Standards or Framework?</td>
<td>d) Full questionnaire/assessment available?</td>
<td>Findings</td>
</tr>
<tr>
<td>--------------------</td>
<td>----------------------------------</td>
<td>-----------</td>
<td>------------------------------------------------------------</td>
<td>--------------------------</td>
<td>----------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Wema, 2006</td>
<td>a) 9 sets of questions, all questions were True/False/No Comment. Sections with number of questions: Defining information problem or research question – 5 Information sources – 10 Internet sources – 8 Internet searching – 12 Library and database searching – 8 Evaluating information and sources – 13 Referencing – 10 Synthesizing information – 6 Presenting information – 8</td>
<td>b) Pre- and post-test were given during the 7-day training session but no mention of exact timing.</td>
<td>c) Recall.</td>
<td>a) Yes, instruments were tested prior to use. Program and assessments were piloted with a group of librarians from the same institution before use in the study.</td>
<td>b) Instrument was based on a questionnaire by Andretta (2005) plus others not specified, with adjustments made to reflect needs of setting and participants.</td>
<td>c) Yes (ACRL IL standards plus IL standards from other countries).</td>
<td>d) Yes.</td>
<td>Students’ scores increased on average about 30 points, but no statistical analysis provided.</td>
<td>Quizzes to encourage reflection &amp; test understanding; assessment of presentations (at the end of each module and on the last day of the program) to determine strengths and weaknesses in applying what was learned.</td>
</tr>
</tbody>
</table>
as a moderator was considered first by comparing the three types of format: face-to-face (nine effect sizes), online (three effect sizes), and hybrid (four effect sizes). Results of the analysis were not significant (QM (df=2) =0.77, \(P=0.68\)) indicating there was only random variation in effect sizes between format types.

Discipline of students was also considered as a potential moderator. Other than medicine and education, no discipline had more than two associated studies, and many had only one (see Figure 2). However, studies were almost equally divided between those involving health science students (seven effect sizes from six studies) and non-health science students (nine effect sizes from six studies), so those two groups were compared. There was a significant difference in effect size based on discipline as a moderator (QM (df=1) =6.54, \(P=0.01\)), therefore, two additional meta-analyses were run. For studies involving only health science students there was a lower SMD of 0.60 [SE=0.17, \(z=3.32, P=0.009\), 95% CI=0.23-0.88] while for non-health science students the SMD increased to 1.43 [SE=0.30, \(z=4.83, P<0.001\), 95% CI=0.85-2.00].

In a model including both moderators (format plus discipline), the test for residual heterogeneity was significant (QE (df=12) = 40.23, \(P<0.001\)) indicating that other moderators, not included in the analysis, are potentially influencing the effectiveness of instruction (Viechtbauer, 2010). Duration of instruction could be expected to influence effectiveness; however, several studies failed to include duration of instruction. When information about duration was provided, length varied widely.
For example, for sessions provided within a subject class, duration ranged from one 70-minute session to two 3-hour sessions. The small number of studies in each duration length precluded completing a subgroup analysis of duration.

**Publication bias**

The funnel plot for the meta-analysis of all RM studies is shown in Figure 4. Studies seem to be evenly distributed at the top of the funnel but lacking toward the bottom. However, the fail-safe number was calculated to be 750, considerably larger than the minimum of 90 suggested by Rosenthal ($5n + 10 = 5(16) + 10$). Since 750 non-significant studies would be required to reduce the overall effect size to zero, publication bias was not considered an issue.

**Discussion**

There was a positive overall SMD, which suggests that library instruction does increase information literacy knowledge and/or skills in graduate students, and that the average increase in score is about one standard deviation. Although this appears to be the first systematic review and meta-analysis involving library instruction for graduate students, there is a previous meta-analysis of library instruction for undergraduates, which found similar results (Koufogiannakis & Wiebe, 2006). Like this study, Koufogiannakis and Wiebe (2006) found a positive effect when comparing library instruction to no instruction, but the effect was much smaller, about one-third of a standard deviation (SMD=0.36, 95% CI=0.14-0.50). The smaller effect may be explained by the fact that Koufogiannakis and Wiebe (2006) were...
comparing only traditional (passive) instruction to no instruction, while this study compared all types of instruction to no instruction. Koufogiannakis and Wiebe (2006) also compared traditional instruction to computer-aided (online) instruction, and like this study, found no difference in the effectiveness of the two formats. However, findings about hybrid instruction from this study differ from those of another meta-analysis of blended (hybrid) learning in health professions (Liu et al., 2016). While Liu et al. (2016) concluded that blended (hybrid) instruction was more effective than non-blended instruction (SMD=0.81, 95% CI=0.57-1.05), this study found that there was no statistically significant difference between effect sizes for different formats of instruction, including hybrid, face-to-face, and online. One difference between the two studies is sample size; the small number of studies involving hybrid instruction in this meta-analysis limits the robustness of those results.

Small numbers of studies also impacted the ability to look at effect of instruction by discipline of students. Two broad categories (health science students and non-health science students) were examined rather than individual disciplines. Findings indicated a significant difference in effect size between instruction for health science and non-health science students, with library instruction for health science students slightly less effective (average increase of about two-thirds of a standard deviation) than library instruction for non-health science students (average increase of almost 1.5 standard deviations). This result may be explained in part by the likelihood that assessing the ability to apply knowledge results in smaller changes than simply testing students’ recall of information. More than 40% of the studies that included health science students assessed application of knowledge. In contrast, the studies involving non-health students all assessed recall of knowledge, although two of them did also include a few questions that required students to apply knowledge in order to answer multiple-choice questions.

Limitations

One limitation for the overall meta-analysis was the lack of required information from studies, resulting in the need to contact authors and if that failed, to estimate standard deviation for some studies. As pointed out by Gerstner et al. (2017), effective meta-analyses rely on complete data reporting in primary studies. To ensure more complete and accurate meta-analysis of results, studies reporting educational interventions with pre- and post-assessments should either include pre- and post-means and standard deviations or provide raw data so that those statistics can be calculated. A second limitation in the subgroup analyses was the small number of studies in some categories. Duration of intervention, which might be expected to affect effectiveness, was not considered for subgroup analysis because of the lack of information in some studies and lack of uniformity of duration in the remaining studies. In addition, when examining format of instruction, there were only three studies involving online instruction and four with hybrid instruction. Borenstein et al. (2009) point out that in a random effects model, small numbers of studies make it more difficult to estimate error and increase the possibility of not only an inaccurate effect, but also an inaccurate range of effect. Therefore, results of any subgroup analysis with a small number of studies must be regarded with caution. Small numbers of studies may have also affected subgroup analysis of instruction by discipline since no discipline had more than four studies, and studies had to be combined into much broader categories of health science and non-health science students.

Implications for Practice

- Library instruction for graduate students seems to be effective in
increasing students’ knowledge and skills.
• There was no significant difference in effectiveness of face-to-face, online, or hybrid formats of instruction.
• Content varied but information about searching effectively was present in all studies.
• Evaluating students’ ability to apply what they learned rather than testing recall of facts may be a more accurate evaluation of instructional impact.
• Most researchers created their own evaluation instrument. Using existing validated instruments would allow more robust comparisons.
• There is a need for more published studies (particularly for non-health science disciplines) and for more complete reporting of study design including information about timing, duration, and content.

Conclusion

In the current climate of accountability in higher education, it is important to know whether the time and effort spent on providing library instruction for graduate students is effective in producing an increase in information literacy knowledge and skills. However, studies involving library instruction often lack power due to small sample sizes; combining studies in a meta-analysis to determine an overall effect size can overcome that problem. This review found 12 repeated measures studies and four independent group studies that tested the impact of library instruction. Meta-analysis of the 12 repeated measures studies indicate that library instruction for graduate students was effective in increasing information literacy knowledge and/or skills on average by about one standard deviation. Subgroup analysis found a significant moderation of effect between two broad categories of health science and non-health science students. Studies involving health science students resulted in a smaller increase of almost two-thirds of a standard deviation, while studies of non-health science students had an increase of almost 1.5 standard deviations. The difference in the two groups may be the result of a difference in assessment, with health science studies more likely to assess application of knowledge rather than recall of information.

Results of subgroup analyses must be viewed with caution due to small numbers of studies in most subgroups. To strengthen the accuracy of future meta-analyses, there is a need for larger numbers of studies that measure the impact of library instruction, particularly instruction provided in an online or hybrid format. There is also a need for precise description of instructional sessions and more robust data reporting by authors of primary studies.

Acknowledgements

The authors acknowledge with grateful appreciation the assistance of Dr. Alan Wilson of the School of Fisheries, Aquaculture and Aquatic Sciences at Auburn University in completing this meta-analysis and manuscript.

References


Koufogiannakis, D. & Wiebe, N. (2006). Effective methods for teaching information literacy skills to undergraduate students: A systematic review and meta-


Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097. [https://doi.org/10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097)


https://doi.org/10.1016/s0895-4356(99)00161-4

Retrieved from http://www.jstatsoft.org/v36/i03/
Appendix A

Search Strategies

All searches were run on 11 March 2019 and were limited to English language and a date range of 2000-2019.

1. The following five databases were searched concurrently through the EBSCO interface with the “Select a Field” option:
   - Library Literature and Information Science Index (H. W. Wilson)
   - Library, Information Science & Technology Abstracts
   - Medline
   - CINAHL
   - ERIC

   Using this search:
   - Librar* AND (Information literacy OR instruct* OR train* OR orient* OR educat* OR library user education OR library instruction OR library orientation) AND (Random* OR RCT OR (“pre” AND post) OR (before AND after)) AND (((Graduate OR masters OR doctoral OR PhD) AND (student* OR study OR studies OR program* OR degree OR education)) OR medical student* OR dental student* OR professional student*)

2. Library and Information Science Abstracts (LISA) was searched through the ProQuest interface using this search:
   - Librar? AND (Information literacy OR instruct? OR train? OR orient? OR educat? OR library user education OR library instruction OR library orientation) AND (Random? OR RCT OR (“pre” AND post) OR (before AND after)) AND (((Graduate OR masters OR doctoral OR PhD) AND (student? OR study OR studies OR program? OR degree OR education)) OR medical student? OR professional student? OR dental student?)

3. ProQuest Dissertations and Theses Global was searched using this search:
   - AB((Librar?) AND (Information literacy OR instruct? OR train? OR orient? OR educat? OR library user education OR library instruction OR library orientation) AND (Random? OR RCT OR (“pre” AND post) OR (before AND after)) AND (((Graduate OR masters OR doctoral OR PhD) AND (student? OR study OR studies OR program? OR degree OR education))) OR TI((Librar?) AND (Information literacy OR instruct OR train? OR orient? OR educat? OR library user education OR library instruction OR library orientation) AND (Random? OR RCT OR (“pre” AND post) OR (before AND after)) AND (((Graduate OR masters OR doctoral OR PhD) AND (student? OR study OR studies OR program? OR degree OR education)) OR medical student? OR professional student? OR dental student? OR professional student?))

1 “Select a field” searches the author, subject, keyword, title, and abstract fields. (More information here: https://help.ebsco.com/interfaces/EBSCO_Guides/General_Product_FAQs/fields_searched_using_Select_a_Field_drop_down_list)
### Appendix B
Data used in Meta-Analyses

<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>M</th>
<th>SD</th>
<th>N</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARONOFF ET AL., 2017</td>
<td>40</td>
<td>64</td>
<td>16</td>
<td>40</td>
<td>73</td>
<td>16</td>
</tr>
<tr>
<td>BEILE ET AL., 2004 GROUP 1</td>
<td>16</td>
<td>60</td>
<td>9.83</td>
<td>16</td>
<td>70.63</td>
<td>11.53</td>
</tr>
<tr>
<td>BEILE ET AL., 2004 GROUP 2</td>
<td>19</td>
<td>54.21</td>
<td>14.65</td>
<td>19</td>
<td>71.32</td>
<td>12</td>
</tr>
<tr>
<td>BEILE ET AL., 2004 GROUP 3</td>
<td>14</td>
<td>63.57</td>
<td>15.62</td>
<td>14</td>
<td>78.57</td>
<td>13.93</td>
</tr>
<tr>
<td>CHIARELLA ET AL., 2014</td>
<td>61</td>
<td>78.9</td>
<td>15.7</td>
<td>61</td>
<td>79.9</td>
<td>15.8</td>
</tr>
<tr>
<td>DORSCH ET AL., 2004</td>
<td>33</td>
<td>56.11</td>
<td>13.14^2</td>
<td>30</td>
<td>64.08</td>
<td>13.14^2</td>
</tr>
<tr>
<td>EMMETT ET AL., 2007</td>
<td>16</td>
<td>47.5</td>
<td>15.6^1</td>
<td>16</td>
<td>74.1</td>
<td>5.03^1</td>
</tr>
<tr>
<td>GRANT ET AL., 2006</td>
<td>13</td>
<td>4.58</td>
<td>1.5</td>
<td>11</td>
<td>6.45</td>
<td>1.46</td>
</tr>
<tr>
<td>IVANITSKAYA ET AL., 2008</td>
<td>14</td>
<td>39</td>
<td>6.31</td>
<td>14</td>
<td>41.36</td>
<td>6.33</td>
</tr>
<tr>
<td>LECHNER, 2005 GROUP 1</td>
<td>17</td>
<td>48.9</td>
<td>13.14^2</td>
<td>17</td>
<td>67</td>
<td>13.14^2</td>
</tr>
<tr>
<td>LECHNER, 2005 GROUP 2</td>
<td>10</td>
<td>54.5</td>
<td>13.14^2</td>
<td>10</td>
<td>62.6</td>
<td>13.14^2</td>
</tr>
<tr>
<td>MARANDA ET AL., 2016</td>
<td>60</td>
<td>43.33</td>
<td>35.01</td>
<td>60</td>
<td>52.5</td>
<td>38.43</td>
</tr>
<tr>
<td>OTTO, 2012</td>
<td>13</td>
<td>5.77</td>
<td>1.74^3</td>
<td>13</td>
<td>7.62</td>
<td>2.1^3</td>
</tr>
<tr>
<td>SHAFFER, 2011 GROUP 2</td>
<td>30</td>
<td>40.83</td>
<td>14.568</td>
<td>30</td>
<td>63.33</td>
<td>16.312</td>
</tr>
<tr>
<td>WEMA, 2006</td>
<td>12</td>
<td>31.8</td>
<td>9.09^3</td>
<td>12</td>
<td>68.7</td>
<td>6.23^3</td>
</tr>
</tbody>
</table>

n = number of participants, m = mean, sd = standard deviation, _pre refers to pre-assessment, _post refers to post-assessment

**Notes:**

1 Information not provided in article; author provided data by email
2 SD not available, used average of other included studies to estimate SD (see Furukawa et al., 2006 for justification)
3 Information not provided in article, calculated from available raw scores
Appendix C

Sample R Code

# meta-analysis using SMDH
> library(metafor)
> data=read.csv(file.choose())
> data
> mymeta<- escalc(measure="SMDH", n1i=n_post, m1i=m_post, sd1i=sd_post, n2i=n_pre, m2i=m_pre, 
sd2i=sd_pre, data=data)
> mymeta
> results=rma(yi, vi, data=mymeta, slab=paste(Study), method="REML")
> results
> forest(results)
> funnel(results)
# effects of moderators
> resm=rma(yi, vi, mods=~factor(allocest),data=mymeta, method="REML")
> resm
> resf<rma(yi, vi, mods=~allocf, data=mymeta)
> resf
> resd<rma(yi, vi, mods=~allocd, data=mymeta)
> resd
# combined effects of moderators to test residual heterogeneity
> res<rma(yi, vi, mods=~allocd + allocf, data=mymeta)
> res
# Rosenthal’s fail-safe number
> fsn (yi, vi, data=mymeta)
Appendix D

Studies Included in the Systematic Review

Note that studies with an * were not included in the meta-analysis.


*Ilic, D., Tepper, K., & Misso, M. (2012). Teaching evidence-based medicine literature searching skills to medical students during the clinical years: A randomized controlled trial. *Journal of the Medical Library Association*, 100(3), 190-196. [https://doi.org/10.3163/1536-5050.100.3.009](https://doi.org/10.3163/1536-5050.100.3.009)


