CHAPTER 6
WHAT THE CITATION PROJECT TELLS US ABOUT INFORMATION LITERACY IN COLLEGE COMPOSITION

Sandra Jamieson
Drew University

INTRODUCTION

In the introduction to the Framework for Information Literacy for Higher Education (Framework for IL) (ACRL, 2015), the ACRL Board explains (through footnotes) that the thinking behind the Framework for IL is indebted to Thomas Mackey and Trudi Jacobson’s (2011, 2014) work on metaliteracy. That work, it notes,

expands the scope of traditional information skills (i.e., determine, access, locate, understand, produce, and use information) to include the collaborative production and sharing of information in participatory digital environments (collaborate, produce, and share). This approach requires an ongoing adaptation to emerging technologies and an understanding of the critical thinking and reflection required to engage in these spaces as producers, collaborators, and distributors (footnote 7, citing Mackey & Jacobson, 2014).

As writing teachers and librarians develop ways to help students acquire the dispositions identified in the Framework for IL, it is useful to look closely at the kind of researched writing produced before its introduction. In addition to providing a sense of the kinds of resources being consulted in response to specific writing contexts, such analysis provides a baseline to work beyond—and instructors against. We have learned a lot about attitudes, practices, and expectations from student interviews such as those by Project Information Literacy (Head & Eisenberg, 2009, 2010; Head, 2013) and protocol analysis (see...
Blackwell-Starnes, Chapter 7, this collection), but analysis of the final product—the research paper—offers insight into how those various habits of mind play out. Review of student research papers produced before the *Framework for IL* reveals why the shift to a metacognitive information literacy (IL) is welcomed by many involved in IL instruction. It also demonstrates the impact of some of the limitations of the ACRL *Information Literacy Competency Standards for Higher Education (IL Standards)* (ACRL, 2000).

As others have noted (see Norgaard & Sinkinson, Chapter 1, this collection), there are many institutional challenges preventing faculty and librarians from working together to develop shared IL pedagogy. Not the least of these is the common location of IL instruction in the required first-year writing course (FYW) where IL assignments and too often product-based, focusing attention away from IL as a process (Blackwell-Starnes, Chapter 7, this collection) and obscuring the vision of IL as “a rich multifaceted literacy that is responsive to changing contexts and opportunities” (Norgaard & Sinkinson, Chapter 1, this collection). Yet the location of IL instruction is not likely to change with the introduction of the new *Framework for IL*, making it essential for those who would develop a more responsive IL pedagogy to understand what happens in this current context. For this reason, analysis of the research papers produced in first-year writing (FYW) prior to 2011 is particularly instructive. FYW is the one college-level course that almost always includes a researched project (Hood, 2010). Whether the instruction takes the form of the “one shot” library visit (Gavin, 1995; Jacobs & Jacobs, 2009), a program-wide IL component (Holliday & Fagerheim, 2006; Jacobson & Mackey, 2007), embedded librarians (Deitering & Jameson, 2008; Kesselman & Watstein, 2009), or team-taught courses (Alvarez & Dimmock, 2007; Jacobson & Mackey, 2007), final papers are expected to reflect what students have learned about research writing (Howard & Jamieson, 2014) and IL. Because IL instruction is often formally or informally assessed based upon those papers, we can also use them to assess the *IL Standards*.

Articles published by librarians and by writing teachers that focus on or include data on papers produced in FYW courses (Grimes & Boening, 2001; Carlson, 2006; McClure & Clink, 2009) or other lower-level introductory courses mostly populated by first-year students (Davis & Cohen, 2001; Jenkins, 2002; Davis, 2002, 2003; Carlson, 2006; Knight-Davis & Sung, 2008) can help us begin this assessment, but all report on single-institution studies, so they may be too limited to allow broad conclusions. Such single-site research has important local relevance, permitting the campus community to explore questions about the kinds of sources selected and retrieved at a specific moment in a specific place and develop responsive pedagogies and policies. Individually,
though, they reveal little about national patterns or trends. These studies occur in isolation and are designed to address local concerns rather than being developed in response to other research. Indeed, single-institution studies are so inward-focused that very few replicate the methods or coding categories of other studies with which they might compare data. Lacking such overlap, existing studies cannot be easily aggregated as part of an evolving national picture of source use. The two exceptions to this are Project Information Literacy and the Citation Project, both of which are engaged in multi-institution study of student research practices and products.

This chapter reports on data from a study of the Citation Project Source-Based Writing (CPSW) Corpus. The study in question explores the types of sources selected and cited in 800 pages of source-based writing by 174 students enrolled in FYW courses at 16 U.S. institutions, ranging from community colleges to Ivy Leagues. Source codes replicate coding categories of earlier studies (Carlson, 2006; McClure & Clink, 2009). Sub-codes allow the data to be broken out for comparison with other studies (Davis & Cohen, 2001; Jenkins, 2002; Davis, 2002, 2003; Knight-Davis & Sung, 2008), extending their reach and reinforcing some of their findings while challenging other oft-repeated claims. This research allows us to understand the limits of decontextualized and linearly focused IL instruction (Norgaard & Sinkinson, Chapter 1, this collection). The data indicate that, nationally, students are broadly able to identify, locate, and access information from apparently appropriate sources in sanctioned ways; however, a closer look at which texts are cited and the ways they are incorporated into the papers reveals the need to go beyond what has for many become a checklist mentality to what the Framework for IL describes as “an expanded definition of information literacy [that] emphasize[s] dynamism, flexibility, individual growth, and community learning” (“Introduction”).

THE CITATION PROJECT DATA

CITATION PROJECT SOURCE-BASED WRITING (CPSW) CORPUS

As reported elsewhere (Jamieson & Howard, 2013; Jamieson, 2013), the Citation Project Source-Based Writing Corpus (CPSW) gathered research papers from 16 institutions distributed regionally and representing 12 states throughout the United States and also distributed across 2008-2010 Carnegie classifications (see Table 6.1). The papers were produced at the end of whatever the institution identified as the standard FYW course, requiring a 7–10-page research paper using at least five sources. Only decontextualized final research papers were collected; Institutional Research Board (IRB) approvals required researchers to
protect students and faculty from possible repercussion should plagiarism be detected, which prevented collection of any demographic information, syllabi, or assignments. The study focuses only on the finished product of the research process—the papers—gathered between Spring 2008 and Spring 2010, and the source use in a total of 50 pages from each participating institution was coded. In all, 174 papers and works cited lists were examined, along with the 1,911 citations they included and the 930 sources cited.

**Locating Sources Cited**

In order to generate the 174 papers used in this study, 171 papers were rejected (Table 6.1) because researchers were unable to retrieve all of the sources listed. In
some cases, the irretrievable sources were part of localized databases set up by the participating institution but not accessible to researchers, or were part of larger collections behind a prohibitive pay-wall. In other cases, citations provided inadequate documentation, especially URLs containing typographical errors; many others pointed to URLs that no longer exist or had been overwritten. Other bibliographic studies spend considerable time discussing concerns about unretrievable Internet sources. Grimes and Boening (2001) note that 30% of the URLs cited in their sample could not be found “due to either student misreporting of the URL or inactive links” (p. 19); and when Davis (2003) checked URLs for “accuracy and persistence” six months after collecting the papers in his 2001 study, he found that 35% did not take him to the original source (p. 55).

In general, these researchers attribute their difficulty locating cited sources to errors on the part of the students or imply that the difficulty reveals the inadequacy of the source itself, citing both as further evidence of the need to strengthen IL instruction. It is possible, though, that many of the URLs were correct when the students listed them. Lepore (2015) notes that “the average life of a webpage is 100 days,” and Zittrain, Albert, and Lessig (2014) found that 70% of the 1,002 sampled URLs cited in the Harvard Law Review, the Harvard Journal of Law and Technology, and the Harvard Human Rights Journal failed to send readers to the information originally cited (what they term “reference rot”). The same was true of 49.9% of the 555 URLs in all published United States Supreme Court opinions. Instead of jumping too quickly to conclusions about the quality of Internet sources selected by the students or blaming student honesty or IL skills for irretrievable sources, these findings suggest that teachers and librarians should revise IL instruction to include discussion of the role of accurate citations, the problem of “reference rot,” and the importance of listing DOIs if they exist.

The fact that, using the information provided, Citation Project researchers were able to locate most of the Internet sources in the sample papers using the Internet Archive (http://archive.org/), suggests that reference rot might have been part of the problem in other studies as well. Successful retrieval of sources by Citation Project researchers was higher when papers used MLA-style works cited lists that note access date, but approximate searches based on the date of the paper were also largely effective. The Internet Archive allowed Citation Project researchers to read Internet sources as they appeared the day the student consulted them.

**Coding Categories**

Data on source use has been reported elsewhere, along with a discussion of methods (Jamieson & Howard, 2013, Jamieson, 2013). This chapter focuses
on the sources themselves, which were classified into one of 14 types (see Table 6.2). The category “book” is uniformly described across studies of student source use (Davis & Cohen, 2001; Jenkins, 2002; Davis, 2002, 2003; Carlson, 2006; Knight-Davis & Sun, 2008; McClure & Clink, 2009); Citation Project coding replicated this category and, like other studies, included books accessed in any format. The definition of journal article as peer-reviewed and written for an academic audience is also quite standard, although Citation Project coding replicated Jake Carlson’s (2006) language, “written by an academic expert in the field, incorporating scholarly perspectives such as theory or research, and having an intended audience of other individuals knowledgeable in the field” (p. 16). The category “Specialized News Source and other periodicals” is defined by Citation Project coding the way Carlson defines “Magazine article” (“reporting an event, opinion, or other issue from a non-scholarly perspective . . . written in a way that would be accessible to a general audience,” p. 16), and includes articles from publications such as The Economist, Nature, Mother Jones, The New Yorker, and Harpers, regardless of how they were accessed. In the case of encyclopedia, dictionaries, and government documents, again no distinction was made between those consulted electronically and those consulted in print, although almost all of the citations indicated that they were consulted online. The category “General News Source” includes traditional newspapers that appear in print and electronically, as well as news delivered by television and radio and related websites (where broadcast news and related information is repeated and updated), and via apps, social media, and email and text updates. Neither the reputation nor the politics of the news source were noted, although sources were also coded using a slight modification of the categories developed by McClure and Clink (2009) as “information (apparently without bias),” “opinion,” “advocacy,” “commercial,” and “self-help.”

While Davis (2003) does note that some websites in his studies would probably be deemed sufficiently informational to be included in student papers, he does not break out URLs in this way, nor do Carlson (2006) or Knight-Davis and Sung (2008). McClure and Clink (2009) do make that distinction, and Citation Project research followed their lead and used the categories they developed. Informational Internet sites are defined as sources that seem to be presenting information without bias or commercial backing, such as the American Cancer Society, and the CDC. Researchers also coded an additional category not included by McClure and Clink, “Internet, multiple-author,” special interest websites or eZines that include articles by a number of contributors but do not have a print version and are not associated with any news or entertainment organizations (most notably fan sites for sports, collectables, or activities). This category includes commercially produced multi-user fandom sites associated
with films, books, or television shows, although such sites were also coded as commercial following McClure and Clink.

Data was analyzed using SPSS (Statistical Package for Social Science 14.0).

LIMITATIONS OF THE STUDY

Although it includes students from 16 institutions, this study only provides a snapshot of first-year college students in the U.S. as they leave a specific course in a given year. As Carlson (2006) notes about his own study, the data are potentially skewed by the fact that the instructors from whose classes the papers were drawn and the students who submitted their papers were volunteers rather than being randomly selected (randomizing occurred after papers had been submitted). Instructors who felt they were doing a good job teaching research skills were probably more likely to volunteer than those who had doubts; experienced instructors were also probably more likely to volunteer. As for the students, those who were misusing sources or who had very low confidence in their research and citation skills probably selected out of the study. Papers were drawn from the standard FYW course, which is sometimes the second writing course for those deemed weaker writers and which in some institutions stronger writers place out of and English Language Learners or multilingual writers are tracked out of. These factors provided a fairly evenly prepared pool of writers for the study, but also limit what might be learned from outliers.

FINDINGS

TYPES OF SOURCES SELECTED AND RETRIEVED

Surveys of faculty expectations for the FYW research paper (see Howard & Jamieson, 2014) and conversations with librarians about sources they recommend indicate that, in general, books, journal articles, government documents, and specialized news sources (accessed either electronically or in print) are considered appropriate sources for FYW research papers. As Head and Eisenberg (2010) among others have found, many assignments still require that students use a specific number of types of sources, often one book and at least two journal articles. Once those requirements are satisfied, many faculty also consider newspaper articles appropriate, although of course this depends on the context and the nature of the topic selected. The majority of FYW research papers address general interest topics selected by the student, and 85% of the assignments studied by Head and Eisenberg (2010) either expected students to generate their own topic or provided acceptable topics from which they could choose (p. 8). This
Table 6.2. Categories of sources selected and used at least once

<table>
<thead>
<tr>
<th>Source Type</th>
<th>Frequency</th>
<th>Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book (single author or anthology)</td>
<td>128</td>
<td>13.76</td>
<td>13.76</td>
</tr>
<tr>
<td>Journal *</td>
<td>219</td>
<td>23.55</td>
<td>37.31</td>
</tr>
<tr>
<td>Specialized news source or periodical *</td>
<td>105</td>
<td>11.29</td>
<td>48.60</td>
</tr>
<tr>
<td>Government document or publication *</td>
<td>63</td>
<td>6.77</td>
<td>55.37</td>
</tr>
<tr>
<td>General news source *</td>
<td>141</td>
<td>15.17</td>
<td>70.54</td>
</tr>
<tr>
<td>Visual images (still and moving)</td>
<td>7</td>
<td>0.75</td>
<td>71.29</td>
</tr>
<tr>
<td>Encyclopedia *</td>
<td>18</td>
<td>1.93</td>
<td>73.22</td>
</tr>
<tr>
<td>Dictionary *</td>
<td>10</td>
<td>1.07</td>
<td>74.29</td>
</tr>
<tr>
<td>Informal print or oral (email, text mess., etc.)</td>
<td>4</td>
<td>0.43</td>
<td>74.72</td>
</tr>
<tr>
<td>Public Internet</td>
<td>235</td>
<td>25.28</td>
<td>100.00</td>
</tr>
<tr>
<td>TOTAL</td>
<td>930</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

* accessed electronically or in print

also appears to describe the papers in the CPSW, whose topics are frequently abortion, gun control, Title IX, global warming, marijuana (legalization/health benefits), and Internet privacy. Some papers focus on literature or discipline-specific topics requiring specialized sources, but these are the minority.

As Table 6.2 shows, source types that fit the category “appropriate for FYW” as described above dominate the list of sources selected, retrieved, and cited at least once in the 800 pages coded. Of the 930 sources cited in coded extracts written by the 174 student participants, 55% fall into this category. Of those, books make up 14%; articles from scholarly journals, 24%; specialized news sources and periodicals, 11%; and government documents, 7%. If general news sources (15%) and visual images, mostly films (0.75%) are added to the list of generally acceptable source types, as they often are in FYW courses, the percentage of the 930 sources that would be considered appropriate types for FYW rises to 71%. Relatively few students cite encyclopedia and dictionaries (2% and 1%), or informal print or oral sources (0.5%), although this does not mean they do not use them; only cited sources were studied.

A Closer Look at Sources

Where sources were available in print or electronically, they were coded by source type (e.g., journal article) rather than method of retrieval, but sources that listed a URL were also coded as “Internet” following previous research. Table 6.3
reveals more about the 235 Internet sources cited at least once within the coded pages. More than half (54%) are informational websites, which is 14% of all of the sources cited at least once. Such informational sources (including the American Cancer Society and the CDC) are generally also acceptable in FYW courses, raising the percentage of the 930 sources that would be considered appropriate types of sources by most writing teachers today to 85%.

Of the 107 Internet sites not classified as “informational,” Table 6.4 shows that 21 (9% of all 235 Internet sites) fit McClure and Clink’s (2009) classification of “advocacy website,” and 7 (3%) are websites with clearly commercial motivation. While many websites are sponsored or include advertising, this

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Percent of all internet sources (n=235)</th>
<th>Percent of all 930 sources (Table 6.2)</th>
<th>Cumulative percent of all sources (n=930)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informational website</td>
<td>128</td>
<td>54.47</td>
<td>13.76</td>
</tr>
<tr>
<td>Personal website incl. social media</td>
<td>5</td>
<td>2.13</td>
<td>0.54</td>
</tr>
<tr>
<td>Blog (personal or professional)</td>
<td>14</td>
<td>5.96</td>
<td>1.50</td>
</tr>
<tr>
<td>Multiple-author (eZine, wiki, etc.)</td>
<td>32</td>
<td>13.62</td>
<td>3.45</td>
</tr>
<tr>
<td>Other (not classified above)</td>
<td>56</td>
<td>23.82</td>
<td>6.03</td>
</tr>
<tr>
<td>TOTAL</td>
<td>235</td>
<td>100.00</td>
<td>25.28</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Percent of all internet sources (n=235)</th>
<th>Percent of all 930 sources (Table 6.2)</th>
<th>Cumulative percent of all sources (n=930)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Informational (no obvious bias)</td>
<td>128</td>
<td>54.47</td>
<td>13.76</td>
</tr>
<tr>
<td>Advocacy</td>
<td>21</td>
<td>8.94</td>
<td>2.26</td>
</tr>
<tr>
<td>Personal</td>
<td>19</td>
<td>8.08</td>
<td>2.04</td>
</tr>
<tr>
<td>Company or commercial</td>
<td>7</td>
<td>2.98</td>
<td>0.75</td>
</tr>
<tr>
<td>Online journal (unsponsored)</td>
<td>32</td>
<td>13.62</td>
<td>3.45</td>
</tr>
<tr>
<td>Other (not classified above)</td>
<td>28</td>
<td>11.91</td>
<td>3.02</td>
</tr>
<tr>
<td>TOTAL</td>
<td>235</td>
<td>100.00</td>
<td>25.28</td>
</tr>
</tbody>
</table>
category focuses on websites whose purpose is directly or indirectly commercial (selling or promoting an item, brand, person, location, activity, etc.). These findings are significantly lower than McClure and Clink’s, revealing again the influence of context on single-institution data.

**A Closer Look at Books**

When the category “book” is sub-divided, the question of acceptable source type is further complicated, as Table 6.5 shows. Of the 128 books selected, retrieved, and used at least once in the 800 coded pages, 49% are the traditional scholarly texts probably imagined as the result of the instruction to “include at least one book.” Such texts make up 7% of the total 930 sources. A further 16% of the 128 books cited are works of literature, and 9% are literacy criticism focusing on them. Of the remaining books, 13% are non-academic (self-help or popular press books that do not cite sources or include notes regarding sources), and 12% are curated collections, such as the *Opposing Viewpoints* and *At Issue* series, and short single-topic textbooks that arrange extracts from longer texts into a “conversation” for the students.

Perhaps unsurprisingly, while some books are cited frequently in literature-based papers, of the 930 sources only 14% are cited more than three times, and 56% are cited only once (Table 6.6).
Table 6.6. Frequency of citation for each of the 930 sources

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Once</td>
<td>525</td>
<td>56.45</td>
</tr>
<tr>
<td>Twice</td>
<td>185</td>
<td>19.89</td>
</tr>
<tr>
<td>Three times</td>
<td>89</td>
<td>9.57</td>
</tr>
<tr>
<td>Four times</td>
<td>48</td>
<td>5.16</td>
</tr>
<tr>
<td>Five times</td>
<td>34</td>
<td>3.66</td>
</tr>
<tr>
<td>Six times</td>
<td>19</td>
<td>2.04</td>
</tr>
<tr>
<td>Seven times</td>
<td>10</td>
<td>1.08</td>
</tr>
<tr>
<td>Eight times</td>
<td>9</td>
<td>.97</td>
</tr>
<tr>
<td>Nine times</td>
<td>4</td>
<td>.43</td>
</tr>
<tr>
<td>Ten or more times</td>
<td>7</td>
<td>.75</td>
</tr>
<tr>
<td>TOTAL</td>
<td>930</td>
<td>100.00</td>
</tr>
</tbody>
</table>

**DISCUSSION**

**IDENTIFYING, LOCATING AND RETRIEVING SOURCES**

Bibliographic coding provides information about the kinds and combinations of sources used in each of the 174 papers and also aggregate data about the 930 sources used within the coded pages (sources not used in the coded pages were not retrieved or coded). It also reveals the frequency of use of each source across the 1,911 citations in the sample, allowing comparison between what was selected for the works cited list and what was actually used by the student to build an argument within the paper itself. This allowed researchers to track what percentage of citations drew on scholarly and non-scholarly sources and also to explore the relationships among the sources cited.

On the face of it, the data in Tables 6.2 to 6.5 indicate that students seem to be able to retrieve *types* of sources that meet faculty requirements and that would probably be recommended by librarians, thereby demonstrating an ability to identify, locate, and access source types appropriate for their academic projects. In spite of the inclusion of obviously non-academic sources and commercial and advocacy websites, Tables 6.2 to 6.5 reveal that the majority of the sources being cited at least once are of a type that most instructors of FYW courses would consider “acceptable.” Tables 6.2 and 6.3 reveal that 85% of the 930 sources (791) would appear to be of acceptable types (although this does not mean that the individual sources would be acceptable to support or build the argument in question). Even when adjusted to remove non-academic books (see Table 6.5), that number is still 80% (774 sources).
While sources classified as “magazines” (such as *Business Week*, *The Economist*, and *National Geographic*) may not be appropriate for research papers in courses like those in economics studied by Davis and Cohen (2001) and Davis (2002, 2003), following McClure and Clink’s (2009) lead, the Citation Project classifies “specialized news sources or periodicals” as generally appropriate for FYW. These differences are, of course, institution-specific (and sometimes instructor-specific), and lacking assignments and handouts, researchers cannot speak to what was acceptable in each case. At the same time, although literature-based and discipline-specific courses skew the data a little, the overall pattern described here persists across sites.

Because they classify sources by type (book, journal) rather than content, the majority of scholars who study student researched-writing focusing on source use or sources cited classify sources that can be accessed freely from the public Internet as “websites” or “web” if a URL is listed with no further subdivision. This means that the category includes everything from self-help blogs to sites that would be considered appropriate for a paper in a FYW course, such as government- or university-sponsored website. Just as the category “book” tends to be considered appropriate without analysis of content, so “web” tends to be considered inappropriate, and “sources from the Internet” are still forbidden or limited by some faculty, particularly beyond the first year. While Davis (2003) does note that some Internet sources would probably be sufficiently informational to be included in first-year economics papers, his study does not break out URLs in this way, nor do those of Carlson (2006) or Knight-Davis and Sung (2008). McClure and Clink (2009), focus most of their attention on the 48% of sources they coded as “websites,” offering a more granular classification and when Citation Project sources are coded using those classifications (Table 6.4) it is obvious that not all Internet sources should be considered unacceptable for source-based papers in FYW courses. The decision to exclude sources found “online” from studies, and still from some classes, seems increasingly limited given the ubiquity of the Internet, the quality of sources available through it, and the growing sophistication of the so-called “digital natives.” Where earlier studies found cause for concern in the types of Internet sources selected, the research reported in this chapter echoes Carlson’s (2006) observation that the quality of sources revealed in single-institution studies do not appear to be generalizable to a national level. The national snapshot provided by the Citation Project suggests less cause for concern about the Internet and also records a high percentage of sources classified as informational (websites from national organizations such as the American Cancer Institute, or from government sponsored sites like the Center for Disease Control).
Perhaps because they were distracted by the Internet, earlier researchers have not focused on the books students select, apparently imagining “books” by definition to be scholarly. In fact, data regarding this category (Table 6.5) suggests the need to revisit the traditional instruction to “include at least one book.” Some of the courses in the sample were literature-focused, producing papers using literary criticism (including SparkNotes and Cliff’s Notes, which are frequently used as criticism) to discuss literary texts. Another subset of courses aid the “research” process by selecting curated collections such as Opposing Viewpoints and At Issue series or single-topic textbooks. While these collections include scholarly sources, they do most of the IL work for students, arranging extracts from longer texts into a “conversation” rather than asking students to conduct research in order to discover possible conversations themselves. Such collections may help students understand the Framework for IL threshold concept Scholarship as Conversation, and many of the papers also draw on other sources that the students may have selected. The use of curated collections in FYW alongside additional student-selected sources is worth further consideration as part of IL pedagogy, especially if students learn to develop source networks from the works cited lists of texts in those collections (Laskin & Haller, Chapter 11, this collection). Books make up only 14% of the 930 sources, and of those, half are monographs and edited collections that are sufficiently scholarly, although a closer look at the books that would not be acceptable is instructive.

Overall, the Citation Project research indicates that in the area of traditional sources (books and journals) and in non-traditional sources (websites), first-year students are mostly able to identify, find, access, and cite sources in ways that would satisfy traditional bibliographic instruction. If the purpose of IL instruction is to help students navigate library databases and stacks along with the Internet, select sources of an appropriate type on a specific topic, and access and cite them correctly, then it seems to have succeeded. Studies of student bibliographies produced in intermediate- and upper-level courses report a high percentage of books and appropriate scholarly journals (Hovde, 2000; Jenkins, 2002; Kraus, 2002; Carlson, 2006; Mill, 2008), confirming that this skill appears to transfer beyond the first year. While Davis (2003) found that students were using more Internet sources, he reports that they did not do so at the expense of traditional sources but as part of an increase in the total number of sources cited (p. 47). Of course, the IL Standards go far beyond simple information retrieval, calling for a sophisticated understanding of sources and listing characteristics of an equally sophisticated “information literate student” within the performance indicators and outcomes. It is on this level that the Citation Project data suggests the IL Standards have been less than successful in transforming both practice and habits of mind. This observation reinforces critiques that the IL Standards
present a set of “neatly packaged skills which result in a successful product” (Norgaard & Sinkinson, Chapter 1, this collection) rather than encouraging a process-oriented pedagogy. Such a recursive, rhetorically based IL pedagogy is, of course, harder to measure using the kinds of checklists so common in FYW and IL assessment, yet it is more in line with the process-oriented pedagogy called for by Writing and Composition theorists, and perhaps a similar focus on process in the Framework for IL will encourage radical pedagogical revision in both areas.

A Deeper Look at Source Use

A closer look at Tables 6.5 and 6.6 suggests that there is still much IL work to be done. The fact that the papers in this study demonstrate an ability to retrieve appropriate types of sources does not mean students can do that work alone. When instructed to include a certain number of books and peer-reviewed articles on their works cited list, most students will comply; however, if they are not part of a brief literature review it is unlikely that a single citation was what the professor intended. Overall, 56% of the 930 sources and 50% of the 128 books were cited only once, although the 21 works of literature were cited with greater frequency. This finding suggests, again, that students need to be exposed to a more sophisticated model of IL that teaches them to retrieve information as needed from types and numbers of sources appropriate to the topic at hand, not to satisfy a decontextualized checklist.

Other studies (Sherrard, 1986; Hull & Rose, 1990; Pecorari, 2003; Howard, Serviss & Rodrigue, 2010) reveal that students working with sources write from sentences within those sources rather than summarizing extended passages, and in this expanded study they were found to do so in 94% of the citations (Jamieson & Howard, 2013; Jamieson 2013). Together, these data suggest that the focus of concern should not be the sources per se but the ways students engage with them and use them to trace connections and create a conversation among them. As Maid and D’Angelo (Chapter 2, this collection) observe, IL is a “contextualized and situated concept,” and by replacing the “prescriptive and de-contextualized set of skills” with a deeper attention to metaliteracy, the new Framework for ILs may begin to address this.

Information Literacy Dispositions and Habits of Mind

Because most of the students used the citation guidelines included in the 7th edition of the MLA Handbook for Writers of Research Papers (2009), which require the inclusion of the medium (print, Web, DVD) and the date a source
was consulted, it is possible to track the order that sources were retrieved and the ways papers were constructed. Jenkins (2002), Davis (2003), and Carlson (2006) all caution that there is no “average bibliography,” and that observation holds for the Citation Project papers as well; there are, though, patterns within the kinds of sources selected and the way those sources are used. A look at papers that use “acceptable source types” reveals an often torturous research process and paper-writing formula that is far from the “set of integrated abilities encompassing the reflective discovery of information, the understanding of how information is produced and valued, and the use of information in creating new knowledge and participating ethically in communities of learning” identified as a goal of the Framework for IL.

One such paper, Z24, while not being typical, is representative of the many struggles revealed in the coded papers, and reflects many of the concerns that indicate the need to move beyond a checklist model of IL. A review of the works cited list (Appendix A) suggests that few of us would find a problem with it unless we looked at the sources themselves. It includes two books, one scholarly journal article, two additional journal/periodical articles, one government website, two informational sources from the Internet, an article from The New York Times, and a website whose title at least would appear appropriate to the paper topic (obesity).

“AUTHORITY IS CONSTRUCTED AND CONTEXTUAL”

A closer look reveals the problems that can result from an over-dependence on source type to assess authority. Neither of the two books cited in this paper is scholarly, the first obviously so from the title—Skinny Bitch. The second, Compulsive Overeating by Judith Peacock, is 64 pages long but may have appeared reliable to the student because s/he used LexisNexis Academic to access the (8-page) chapter “Who is at Risk?” cited in the paper. The publisher, Life Matters, explain “each book defines the problem, describes its effects, discusses dilemmas teens may face, and provides steps teens can take to move ahead,” giving it a Citation Project classification of “book-self help.” It could have been a useful source for a paper analyzing the kinds of advice given to teens suffering from eating disorders; however, that was not the focus of the paper, and it appears that the student did not assess the contextual nature of the data in this source.

One other source was also retrieved from LexisNexis Academic. Listed as being from Biotech Business Week, the title is “Salad Bars in Every School: United Fresh Applauds New Child Nutrition Bill,” and no author is included by the student, although it is actually listed as the United Fresh Produce Association. Biotech Business Week notes on its “about” page that it publishes “News and
information from pharmaceutical and biotechnology companies, with a focus on business trends and analysis.” The article reports on United Fresh’s president’s visit to Capital Hill to thank legislators for including fresh produce in the New Child Nutrition Bill, crediting Congressman Farr and United Fresh lobbying for this legislation. The student introduces unattributed data about the importance of fresh produce from the article with “research shows. . . .” This source should have raised flags about authority when the student found it (Citation Project researchers classified it as “company or commercial”). A student with a deeper understanding of IL might have considered the role of context and purpose and rejected it.

While Z24 is an outlier in some ways, it reflects the impact of findings by other researchers that students tend to trust the authority of all sources they find “through the library” (Tolar-Burton & Chadwick, 2000). Head and Eisenberg (2009) found that 84% of students reported that scholarly research databases were the library resource they used most frequently (p. 22) and 78% reported that such databases “contain more credible content than the internet” (p. 27). Four of the sources in Z24 were retrieved via Academic Search Complete, which notes that it “provides complete coverage of multidisciplinary academic journals . . . [and] supports high-level research in the key areas of academic study by providing peer-reviewed journals, full-text periodicals, reports, books, and more” (ebscohost). Two others came from The LexisNexis-Academic “about” page notes that it allows researchers to “quickly and easily search full-text documents from over 15,000 credible sources of information and pinpoint relevant information for a wide range of academic research projects . . . [from] comprehensive, authoritative news content, including current coverage and deep archives” (LexisNexis). A final source was retrieved from a third “database,” about.com (“the largest source for Expert content on the Internet that helps users answer questions, solve problems, learn something new or find inspiration” according to the “about” page). In fact, about.com led the student to a government document published by the Center for Nutrition Policy and Promotion, which is part of the USDA, but the student does not even mention that this is government research. Do students know the difference between “peer-reviewed,” “credible,” and “expert”? This student appears not to have, nor to have been able to make any judgment call about these different databases when looking at the sources themselves or seen the difference between a government document and a company-sponsored promotion.

The student does introduce Skinny Bitch as “New York Times Best Selling book,” but aside from that does not appear to consider the authority of the sources selected, and only cites one 23-page chapter entitled “Have no Faith: Government Agencies Don’t Give a Shit About Your Health.” (While this might
have been used in dialogue with the government data, it was not.) Source assessment has always been part of IL instruction and is central to the IL Standards, but as Head and Eisenberg’s (2010) data show, if it is not accompanied by a larger metacognitive discussion of the role of research and what we hope students will learn through the process of researching a topic, they are unlikely to incorporate such assessment into their normal research habits or the way they think about information in general.

**Information Creation as a Process and Information Has Value**

Each of the sources in Z24 is used to introduce a reason why people might be obese (eating disorders, high cost of healthy food, low cost of burgers, lack of government concern), but the very different kinds of sources and the lack of acknowledgement of their context or purpose undermines the argument. While some of the sources selected are appropriate in terms of content and type, the remaining sources are isolated voices on the general topic of the student’s paper, and the information in them is created and presented for a very different audience and purpose. One such source is a self-help blog, “Eat without Guilt,” which promises to help readers “make peace with food, your body, and your weight.” A second comes from The New York Times Well blog and is entitled “A High Price for Healthy Food,” and a third is the book chapter “Compulsive Overeating,” which is cited twice in the paper, with both quotations taken from the same page. Blogs and self-help sources would be appropriate for some kinds of paper (an analysis of the rhetoric of “help,” or the kinds of resources available for people with eating disorders), but to mix self-help with academic research in a paper whose stated aim is to explore the causes of obesity is a questionable decision, and to do so without acknowledgment of the different context or recognition of the different scholarly weight assigned to each, reveals a student who still needs to develop IL skills. If students are able to “assess the fit between an information product’s creation process and a particular information need” they are also on the way to developing “an understanding that their choices impact the purposes for which the information product will be used and the message it conveys” (Framework for IL). Such an awareness may have guided this student to different source choices.

**Searching as Strategic Exploration and Scholarship as Conversation**

Perhaps more interesting than the student’s inability to evaluate the sources s/he cites is the story about the process of construction of the paper that the works
cited list suggests. The student accessed four sources on 5 April, 2010, all four of them via Academic Search Complete. Two come from academic journals (Pediatrics and New Scientist), and the other two are from the Proceedings of the National Academy of Sciences of the United States of America and the Tufts University and Nutrition Letter. The New Scientist article, introduced by the student as “a recent study by New Scientist” is a book review of Supersize Me. If one looks at these four sources, a story appears to be emerging, moving from comments in Supersize Me about corn to “Corn Content of French Fry Oil from National Chain vs. Small Business Restaurants” (Proceedings), and from there to “Who’s Losing the Burger Battle?” (Tufts) and “Nutrition Labeling May Lead to Lower-Calorie Restaurant Meal Choices for Children” (Pediatrics). All of these sources were published in 2010 making it impossible for them to cite or reference each other; however, all seem appropriate to begin an exploration of the topic of obesity, and all are in some ways part of a larger dialogue on the given topic. It would appear that on April 5, the student spent some quality time with a reliable database beginning the process of strategically exploring an evolving topic and entering an ongoing conversation on that topic. Perhaps the class visited the library to start their research or a librarian came to the classroom; certainly someone introduced the student to this database and perhaps helped focus the research question. So far, so good.

But the next reported access of source material is not for another three weeks. On April 25 the Biotech article on salad bars was added to the list, appearing to continue the conversation begun on April 5; however, by April 27 when the four remaining sources were accessed the paper seems to be on the way to its final disconnected list of possible causes of obesity, one per paragraph, each supported by a different source. Two of the sources accessed on April 5 (and still listed as “works cited”) do not appear in the final paper (the Proceedings article and the Pediatrics article). Instead, on April 27 the student lists accessing the blogs “Eat without Guilt” and the The New York Times Well blog, the book chapter “Compulsive Overeating,” and the report on fruit and vegetables from the USDA. The book Skinny Bitch is listed as having been read in print so has no access date. It is cited four times in the paper in two different paragraphs drawing from 4 pages in one 23-page chapter. None of the authors of these sources is directly cited by others, none of the student sources is drawn from the works cited lists of other sources, and no citations from one source appear in any of the other sources.

Because the sources selected in late April do not explicitly respond to other sources about obesity or healthy living, it is easier for them to be used to provide “evidence” for one item on the list of causes of obesity that ultimately organize the paper. Perhaps the need for a list-like outline came externally, but had the student been reading the sources as part of a conversation on the topic of
healthy eating and explicitly creating source networks (Laskin & Haller, Chapter 11, this collection), paper Z24 might have evolved very differently. The student concludes by stating agreement with the author of *Skinny Bitch* that ultimately Americans are responsible for the “obesity epidemic” by remaining naive and ill-informed, but this claim contradicts many of the “causes” that organize the paper. The student neither counters that lack of information by adding to the conversation about obesity, nor makes the argument expressed in the conclusion.

It is this concept of research, and academic sources, as part of an ongoing conversation that seems most missing from the papers in the Citation Project study. The organization of Z24 with one source per paragraph (and one paragraph per source) makes it typical. When each source is a discrete item, and even a “type” to be checked off, the possibility of those sources entering into conversation is slim. If assignments do not specify why we do research or how research questions and conversations evolve as Head (2010) found to be the case, and if IL instruction emphasizes identifying, finding, and assessing discrete sources rather than developing metacritical frameworks for thinking about research (Kleinfeld, 2011), it should be no surprise that the resulting papers tend to be “information dumps” rather than forays into academic conversation and the intellectual work the research paper is imagined to be (Robinson & Schlegl, 2004). Diametrically opposite to the image of scholarship as a conversation is the model of research as formulaic, demanding particular types of sources and “killer quotes,” which can mostly be extracted from the first page of the source. This latter version of “The Research Paper” characterizes the papers in the CPSW corpus.

**CONCLUSION**

The data generated by the Citation Project’s multi-institution research suggests that students who received IL instruction in the era of the *IL Standards* have adopted a limited checklist mode of research rather than the nuanced appreciation of sources and source selection included in that document. This suggests that the *IL Standards* themselves did not totally displace the kind of bibliographic instruction that preceded them in 2000. The papers in the CPSW corpus demonstrate that students have the ability to select, retrieve, and cite the right kinds of sources, many of them appropriately academic for first-year papers; however, the papers evidence little or no relationship among those sources. Analysis of the incorporation of information from the sources into the papers reveals students working at the sentence-level, and segregating each source into one paragraph, mostly by including quotation or paraphrase rather than by summarizing larger ideas in a text or comparing arguments with those of other sources. Furthermore, most of the references to sources are to the first
one or two pages and most of the sources are cited only once or twice, suggesting little engagement with them as part of a broader scholarly conversation.

It may be, as I have argued elsewhere (Howard & Jamieson, 2014), that the first-year research paper itself bears more responsibility for this state of affairs than the IL instruction that tries to support it. It is certainly the case that shorter source-based papers could be used to introduce students to IL practices and habits of mind; however, it is also the case that the introduction of threshold concepts in writing pedagogy (Adler-Kassner, Majewski & Koshnick, 2012) and in the Framework for IL increases the likelihood that students may gain a “vision of information literacy as an overarching set of abilities in which students are consumers and creators of information who can participate successfully in collaborative spaces” (“Introduction”). Had the students writing the papers studied by the Citation Project evidenced the kinds of critical self-reflection described in the Framework for IL and embedded in the description of what constitutes IL, it is difficult to believe they could have produced the papers in the sample. We might hope that with the revised instruction associated with threshold concepts, future papers will show evidence of the “dynamism, flexibility, individual growth, and community learning” described in the Framework for IL document (“Introduction”).

This optimism is unlikely to bear fruit unless we ask some difficult questions of IL instruction and the FYW courses where it so frequently becomes ghettoized. What is the role of real IL (not bibliographic instruction) in FYW? How can we ensure that the skills and habits of mind are transferrable to other courses (and to work, life, etc.)? When included amongst many other elements of a writing course, how can threshold concepts be introduced without overwhelming the students—or the course? And if institutions recognize that IL cannot be “delivered” in one library visit, assignment, or even semester, how can it be advanced programmatically or throughout a student’s education (and beyond to lifelong learning)? Finally, how can the Framework for IL be introduced to all faculty—including library faculty, administrators, and students in a way that will help us all to recognize our shared responsibility for IL and our shared stake in successful IL pedagogy? Unless these questions are addressed, I fear that the 2019 FYW research papers will not look significantly different from those produced in 1999 or 2010.

REFERENCES


**APPENDIX: WORKS CITED LIST FOR PAPER Z24 (AS SUBMITTED)**

Citation Project

**Z24**

**Works Cited**
