Student Constructions of Authority in the Framework Era: A Bibliometric Pilot Study Using a Faceted Taxonomy

James W. Rosenzweig, Mary Thill, and Frank Lambert

Using bibliometric data and a faceted taxonomy first published by Leeder, Markey, and Yakel, this pilot study examines student constructions of authority generated from a sample of 60 research papers by students in a freshman-level English composition course. The taxonomy classifies each source using subfacet attributes of author identity, editorial process, and publication purpose facets that, in combination, provide insight into how students navigate today's information ecosystem. The findings suggest that students use a similar array of sources regardless of their demographic background or their academic ability and that the characteristics of these sources have important implications for information literacy instruction and collection development.

Introduction

Librarians' concerns with source evaluation and critical thinking have intensified recently with the rise of a new "post-truth" era. Even as we encourage students to consult the expanding body of high-quality journalistic and scholarly sources available on the open web, students performing internet-based research risk exposure to false news and predatory journals. Further complicating matters, partisan propaganda urging the public to mistrust mainstream media and scholarship exists alongside thoughtful criticisms of biased journalistic and academic systems. With so many competing arguments disseminated each day, educators need to promote effective and practical methods for evaluating information, particularly to novice researchers.

The *Framework for Information Literacy for Higher Education* responds to the current information zeitgeist by pivoting from information literacy standards to more complex information literacy concepts, or frames. The *Framework's* broad guidelines for teaching source evaluation fall under the frame that "authority is constructed and contextual."[1] The *Framework* recommends that novice researchers use "basic indicators of authority, such as type of publication or author credentials" to determine whether a source is trustworthy, while remembering that the systems that create and elevate certain publication types and author credentials are fal-

*James W. Rosenzweig is Education Librarian at Eastern Washington University; email: jrosenzweig@ewu.edu. Mary Thill is Reference Coordinator and Humanities Librarian at Northeastern Illinois University; email: m-thill@neiu.edu. Frank Lambert is an Assistant Professor and MLS Program Director in the College of Education at Middle Tennessee State University; email: Frank.Lambert@mtsu.edu. ©2019 James W. Rosenzweig, Mary Thill, and Frank Lambert, Attribution-NonCommercial (http://creativecommons.org/licenses/by-nc/4.0/) CC BY-NC.

lible.[2] The *Framework* asserts that whether or not a source can be deemed authoritative derives largely from the context in which the information was produced and how it is to be applied.

This pilot project began with two practicing academic librarians and former colleagues who shared the goal of developing a *Framework*-friendly method for understanding how novice researchers in today's uncertain information ecosystem construct authority. This team was joined in mid-2017 by a professor of library and information science who has experience with bibliometric methodologies and statistical analysis. As information professionals with sophisticated search strategies and advanced evaluation criteria, the members of the research team fall closer to the category of the "expert researcher" identified in the *Framework*. Therefore, we do not naturally view the information landscape from the perspective of the novice researcher. However, to adapt our pedagogy effectively as educators, we believe that the profession needs to understand how our students select sources and what sources they consider trustworthy.

The research team used data collected from a sample of 60 research papers and bibliographies from students enrolled in English 102 at a midsized public university in the Midwest. We performed a broad review of the literature to identify a suitable method for analyzing the authority of student sources, eventually settling on and making modest modifications to a faceted taxonomy developed by Chris Leeder, Karen Markey, and Elizabeth Yakel in 2012.[3] The taxonomy of Leeder, Markey, and Yakel is a "format neutral" rating system that eliminates artificial distinctions between print and online resources. Moreover, the taxonomy defines authority contextually, providing librarians and students alike with a method to describe today's complicated scholarly and nonscholarly information landscape with greater precision.

To avoid social desirability bias and to provide the most authentic portrait of student research, we opted to examine student constructions of authority with direct measures (in other words, the authority of the resources students themselves selected to list and cite within papers submitted for a grade), as opposed to self-reported measures (such as interviews with students on what students consider authoritative). We were less interested in students' motivations than we were in understanding and assessing the kinds of authorities present in the English 102 student information landscape as a whole.

This research provides information literacy professionals with a proposed model for assessing lower-level undergraduate student bibliographies within the *Framework*. The categories established by the taxonomy can be useful in clarifying the *Framework* for campus stakeholders. Moreover, the results contribute to the literature that describes the information resources used currently by composition students in their research papers. This application of the taxonomy spotlights potentially problematic source categories, many of which seem likely to proliferate in the current online information environment.

Literature Review

To understand today's complex information environment, we surveyed the literature to identify a method of classifying sources that could equitably address both print and online materials. As many previous studies were designed to measure the use of library-based resources, few offered meaningful methods to describe the varieties of sources now available on the open web.

The emphasis on library materials has led researchers to classify materials from the open web in broad and occasionally reductive fashion. Websites have frequently been grouped together in a single monolithic category, and generally treated as "nonscholarly."[4] While some researchers have classified web sources by top-level domains (for instance, .com, .org, .edu),

this method, by their own admission, does not account for how ill-defined and amorphous those domains prove to be in practice.[5]

Other researchers have employed a qualitative and standards-based approach to source analysis, including librarian-determined measures of source trustworthiness, quality, bias, vetting, or "scholarliness."[6] While these methods have some useful applications and can demonstrate a high level of interrater reliability, they also are susceptible to the shared biases or prejudices we might have about sources as librarians and trained information professionals.

A final group of studies attempted to establish simpler, more value-neutral criteria for classifying and labeling websites based on features like the establishment of author credentials or an indication of the website's intended purpose. An early example of such a study was conducted by Robinson and Schlegl, who modified Philip Davis's content analysis checklist for print materials for use in website evaluation.[7] Stephanie Rosenblatt later employed Robinson and Schlegl's criteria to measure students' use of scholarly sources subsequent to library instruction.[8] One potential drawback to checklists and similar faceted methods of source analysis is that they may lack the nuance of some qualitative measures. However, the research team believes that the use of a faceted approach would increase the likelihood of findings that might surprise or challenge librarians' assumptions, thus allowing the research team to see the information environment in new ways. Therefore, our hope was to find in the literature a checklist or faceted taxonomy approach that could aid in identifying the indicators of authority envisioned by the *Framework*, without relying excessively on the traditional notions of authority that learners are encouraged to treat with healthy skepticism.

Ultimately, the most useful model we found in the literature was the faceted taxonomy developed and published by Leeder, Markey, and Yakel, which they created specifically in the context of their BiblioBouts library research game.[9] The taxonomy is unusually robust in that it classifies sources across any format or medium according to the same simple, value-neutral criteria. It uses five facets to encode details about sources that collectively allow for the calculation of a source quality rating. In her review of the literature on assessing source quality, Helen Georgas described the Leeder, Markey, and Yakel taxonomy as "the most detailed and structured, and therefore the least subjective" of the published rating scales.[10] Although the taxonomy was originally developed for the purpose of testing the effects of a research game, it has already proven flexible enough to be used in at least one other setting. Recently, Dahlen and Hanson used the taxonomy as a measure of the authority of sources retrieved by students engaged in the use of library search engines, and we concur with their judgment that it is a well-designed tool for assessing the authority of sources.[11]

Given the tool's potential, we identified an important gap in the literature that we intend this pilot study to address. The taxonomy has thus far been applied to narrowly defined contexts: Leeder, Markey, and Yakel used it to examine the sources used on a paper in a specialized undergraduate course, in which all students wrote on the same research topic. Dahlen and Hanson used it to assess student performance on artificial, predesigned research tasks. We were interested in applying the taxonomy more broadly, in the context of a composition class in which students were free to select virtually any topic they desired so they could understand better the contours of the information landscape treated as authoritative by college undergraduates. The other significant deviation made by our pilot study was to elect not to produce numerical scores to indicate each source's quality, as the previous studies have done. While quantitative scoring was useful in answering research questions posed by both Leeder

et al. and Dahlen and Hanson, we were interested in taking a descriptive approach to student sources as opposed to a prescriptive approach.

Checklists that yield a raw numerical score have value in many educational settings, but, given the *Framework*'s emphasis on treating information literacy not as a set of skills in which to meet a standard but rather as a set of evolving dispositions that are constantly refined in the mind of the learner, it was important to us to test a model that did not treat source authority as a quantity that can be measured to an exact value. Rather, it was our aim to test a model that, if successful, could yield useful and insightful descriptions of the characteristics of sources without predetermining the value of those sources. By providing simple language for describing author experience, editorial process, and the purpose of publication, the taxonomy of Leeder, Markey, and Yakel makes more visible the combination of elements that, collectively, allow a learner to construct authority in the manner described by the *Framework.* If that taxonomy could be applied successfully to sources used in papers composed on a wide variety of subjects, as in a freshman composition course, it has the potential to equip educators—both librarians and classroom instructors—with vocabulary that could be shared with students and consequently used by those same students in self-aware and self-critical ways to analyze their own processes for ascribing authority to sources.

Methodology

For this pilot study, the research team worked with the first-year writing program at a public, master's level, Hispanic Serving Institution (HSI) in the Midwestern United States during the Spring 2014 semester. All papers were collected from students who volunteered to participate in this study and were enrolled in English 102, which is part of the university's first-year writing program and run by the English department. As a requirement for select majors, English 102 has a large enrollment and is taught in multiple sections by numerous faculty. The English department requires certain standard elements in English 102, such as learning outcomes, conference policies, and the quantity of writing. The department also requires a specific library instruction curriculum:

> All English 102 sections must schedule a librarian to teach "Introduction to Database Searching" (lesson #4 on the following website), as well as at least one additional lesson chosen by the instructor from this site: https://sites.google.com/site/libraryinstructionmenu/home.

Despite this standardization, English 102 faculty have great liberty to structure their sections accordingly. Instructor-selected course topics span a wide range, from science to popular culture. In most cases, instructors allow students to choose their own subject for research. As a result, students frequently write outside their intended major, exploring topics that are not traditionally academic.

Members of the research team visited individual sections of English 102 to recruit students and obtain signed, IRB-approved consent forms. Volunteer participants provided the research team with a copy of their final research paper and bibliography, as well as permission to access their final paper grade and demographic information. We collected 139 papers from a total possible population of 181 study participants from across 12 sections of English 102 for the project. All papers were anonymized prior to scanning to ensure anonymity and

then were converted into Adobe PDF format to allow for easier analysis by the three research-ers. Cumulative GPA, age, gender, and race/ethnicity were collected from Banner, the uni-versity's enterprise resource planning system, to record basic demographic attributes of the student authors of these papers. A randomly selected sample of 60 scoreable papers (papers that included a bibliography with identifiable citations present in the paper's text), ordered alphabetically by student surname, was used for analysis.

The instrument used to code bibliographies was adapted from the taxonomy of Leeder, Markey, and Yakel (tables 1, 2, and 3, below). We identified three of their taxonomy's facets in particular—Author Identity, Editorial Process, and Publication Purpose—as being very well suited to address the contextual nature of authority in a manner that is consistent with the language and intention of the *Framework for Information Literacy*. We chose to omit subjective category scoring calculations meant to differentiate high-quality and low-quality sources, as we intended to con-sider sources outside the paradigms generally used by librarians and composition instructors, which are enshrined in the numerical values assigned by this taxonomy for scoring purposes. We also chose to set aside the two remaining facets, which dealt with source format and genre, due to our misgivings about the limitations of rigidly format-based analysis seen in previous studies, as well as the *Framework*'s insistence that authoritative content "may include sources of all media

TABLE 1
Author Identity/Facet 3

Author Identity Subfacets	Brief Description
3A: Unknown Authorship	The source indicates that the piece was anonymously written.
3B: Layman	A person without demonstrated expertise in the area being written about.
3C: Corporate Authorship	No single author identified on a work issued by an organization.
3D: Professional-Amateur	A person with a degree in another field, but demonstrating interest, dedication, and experience in the area being written about.
3E: Applied Professional	A person with relevant experience, training, or credentials relevant to the area being written about (such as a journalist with a journalism degree OR substantive professional experience).
3F: Academic Professional	A person with a master's or doctoral degree in the area being written about, which they held at the time the content was published.
3Z: Source Unknown	No information on the category could be found.

TABLE 2
Editorial Process/Facet 4

Editorial Process Subfacets	Brief Description
4A: Self-published	Material made public directly by the author.
4B: Vanity Press	Material the author paid to publish, generally as self-promotion.
4C: Collaborative Editing	Material that is reviewed or edited by multiple (possibly anonymous) collaborators.
4D: Moderated Submissions	Contributed content that has been accepted or approved by someone other than the author but that has not undergone editorial review.
4E: Editor and Editorial Staff	Professionally reviewed and approved by editor/editorial staff.
4F: Peer-reviewed	Professionally reviewed and approved by peer scholars.
4Z: Source Unknown	No information on the category could be found.

TABLE 3	
Publication Purpose/Facet 5	
Publication Purpose Subfacets	**Brief Description**
5A: Personal	Material is published without commercial aims.
5B: Commercial	Material is published for commercial gain.
5C: Nonprofit	Material is published by a nonprofit organization.
5D: K–12 Education	Material is published for educational purposes.
5E: Government	Material is published by the government.
5F: Higher Education	Material is published for an academic audience.
5Z: Source Unknown	No information on the category could be found.

types" and that "authority may be conferred or manifested in unexpected ways."[12] The research team left the subfacet categories from the original taxonomy mostly unaltered, adding only a Category Z to designate sources that the research team was unable to verify, either because of an incomplete student citation or because an online source had ceased to exist, including sources that might be captured for posterity in online tools such as the Internet Archive (https://archive.org/).

The 60 student papers yielded 692 bibliographic references. Two members of our team researched each reference independently to determine whether the student's original source could be located. This process began with attempting to view every source or, in the case of library-provided resources, to view the catalog or database records for each source. In re-searching material online, the Internet Archive's Wayback Machine, a searchable repository of websites, proved especially valuable, as it allowed us to view websites as students had seen them in 2014. In spite of a three-year delay from data collection to final data analysis, we were able to locate 98.9 percent of the referenced sources.

The next step in the process was to examine the 692 sources for claims of author identity, editorial review, and publication purpose. As many online sources did not provide this information adequately, we were obliged to perform background research to establish these characteristics for each source. In performing this background research, we relied exclusively on resources freely available to students of English 102 at the institution, such as Academia.edu, *ProQuest Dissertations and Theses*, and *Ulrich's Directory of Periodicals*, to name a few, to seek clarification. Just as students might do, the research team generally accepted published claims as truth, even when those claims were self-published (such as information in the "about us" portion of a website, or a work history posted to LinkedIn). The rationale for this uncritical acceptance was our intent to consider the information landscape as it presents itself to students. The larger questions of how to critically evaluate and verify credentials online are important and related to this work—but beyond the scope of this study, which aimed for a method of analyzing bibliographies that could be applied both consistently and expediently. The reviewers met regularly by videoconference to discuss each citation before reaching 100 percent agreement in order to discuss the interpretation and application of the taxonomy and to develop consistent rules of interpretation that were recorded as scope notes to be used in subsequent coding (see appendix A).

Once all citations had been categorized, we performed a citation count to establish how heavily students relied on the different source types. When a student cites a source to support an argument, the student is putting forward that source as an authority on that topic. Accordingly, the research team determined that citation analysis, a source analysis method common in library literature, is conducive to comprehending student constructions of author-

ity. To minimize error, the counts were performed synchronously through videoconference. As students of English 102 are often novice writers still familiarizing themselves with the mechanics of paraphrasing, quoting, and MLA formatting, the reviewers needed a simple and consistent method to determine what counted as a citation. It was decided to take the imperfect but clear measure of explicit, parenthetical in-text citations.

To validate the coding of the three facets, two members of the research team did the initial coding, as described above. After a brief norming session with three papers, the third member coded a subset of the original sample, using the interpretation guide (see appendix A), and he then performed an independent count of the number of times each reference was cited in the paper. Once this validating coding was complete, Krippendorff's *alpha* was used to measure the degree of agreement between the original two coders' efforts and that of the third coder. In all cases, $a > 0.70$, surpassing the accepted cutoff for rate of agreement between coders.[13] However, different facets had higher rates of agreement (facet 3, $a = 0.8195$; facet 4, $a = 0.7152$; and facet 5, $a = 0.7336$). This variance indicates that there may be future challenges to coding bibliographies using the current facet-coding scheme we have designed, necessitating further changes to the interpretation guide and additional training for future coders.

Besides reporting measures of central tendency in our findings reported below, we also report the results of both parametric and nonparametric inferential statistical tests used for our analyses. Parametric tests were used for normally distributed ratio and interval variables, and nonparametric tests were used for nonnormally distributed ratio and interval data and for all nominal and ordinal variables.

Findings

The students whose papers were chosen randomly for analysis displayed the following demographic attributes. The majority of student authors were female ($n = 35$, 58.3%), a relatively close representation of the population ($n = 111$, 61.3%) as were most of the other demographic variables featured here. All students in the sample had a mean age of 20.75 and were widely distributed ($sd = 3.616$) with a heavy positive skew, indicating that some considerably older students were enrolled in this class. In fact, 20 percent of the students included in our sample were 22 years and older (ages ranged from 22 to 42), while the remaining 80 percent of students were aged 18–21, inclusive. In terms of where these students were in their respective programs of study, 61.6 percent were freshmen, 35 percent were sophomores, and 0.017 percent

TABLE 4
Ethnic Origin of Students in Sample

Race/Ethnicity	Frequency	Valid Percent	Cumulative Percent
African American/Black	1	1.7	1.7
Asian	12	20.0	21.7
Caucasian	15	25.0	46.7
Hawaiian Pacific	1	1.7	48.3
Hispanic, Any Race	28	46.7	95.0
Native American	1	1.7	96.7
Nonresident Alien	1	1.7	98.3
Unknown	1	1.7	100.0
Total	60	100.0	

each were juniors and seniors. Thus, only juniors and seniors were underrepresented in our sample. In terms of ethnic origin, table 4 shows the ethnicity of the students from our sample, which, again, is a very close representation of our population.

These same students in our sample had an average cumulative GPA of 2.9633 (sd = .65511), which is just slightly higher than our population cumulative GPA (2.83). There is no notable statistical correlation between student age and cumulative GPA in this sample, nor is there any notable statistical correlation between a student's gender and ethnic origin and his/her cumulative GPA. While we make no assumptions that there should be such relationships, a larger sample may reveal such trends.

TABLE 5
Breakdown of Author Identity/Facet 3 Subfacet *References*

Subfacet	Translation	Frequency	Valid Percent	Cumulative Percent
3F	Academic Professional	204	29.5	29.5
3C	Corporate Authorship	203	29.3	58.8
3E	Applied Professional	154	22.3	81.1
3B	Layman	71	10.3	91.3
3D	Professional-Amateur	33	4.8	96.1
3A	Unknown Authorship	16	2.3	98.4
3Z	Source Unknown	11	1.6	100.0
Total		692	100.0	

TABLE 6
Breakdown of Editorial Process/Facet 4 Subfacet *References*

Subfacet	Translation	Frequency	Valid Percent	Cumulative Percent
4E	Editor and Editorial Staff	325	47.0	47.0
4A	Self-published	165	23.8	70.8
4F	Peer-reviewed	125	18.1	88.9
4D	Moderated Submissions	59	8.5	97.4
4Z	Source Unknown	11	1.6	99.0
4C	Collaborative Editing	5	.7	99.7
4B	Vanity Press	2	.3	100.0
Total		692	100.0	

TABLE 7
Breakdown of Publication Purpose/Facet 5 Subfacet *References*

Subfacet	Translation	Frequency	Valid Percent	Cumulative Percent
5B	Commercial	287	41.5	41.5
5F	Higher Education	190	27.5	68.9
5C	Nonprofit	118	17.1	86.0
5E	Government	52	7.5	93.5
5A	Personal	24	3.5	97.0
5Z	Source Unknown	11	1.6	98.6
5D	K–12 Education	10	1.4	100.0
Total		692	100.0	

The application of the taxonomy classified each of the references appearing in each research paper's bibliography according to the three chosen facets. The most frequently occurring subfacets as found in the papers' bibliographies were 3F (Academic professional), 4E (Editor and editorial staff), and 5B (Commercial publication), respectively. The breakdown of each facet by subfacet may be found in tables 5, 6, and 7.

There were 70 combinations of the three combined subfacets (such as $3x4x5x$) coded in our data out of a theoretical 343 possible combinations ($n = 7^3$). The subfacet combination of 3F4F5F (a work written by an academic professional that was peer reviewed and published for higher education purposes) appeared most frequently in the in-text citations of our sample. The subfacet combination 3E4E5B (Applied professional; Editor and editorial staff; Commercial) appeared most commonly in the English composition papers' bibliographies as references. Tables 8 and 9 show the most frequently *cited* subfacet combinations and most frequent subfacet combination *references*, respectively, representing the top 75 percent of combinations. Only five source types, as distinguished and defined by the subfacet combination, out of a possible 343 subfacet combinations account for at least 50 percent of the sources used in terms of being cited and consequently referenced in the papers' bibliographies. Additionally, only 12 source types in table 8, and 14 source types in table 9, account for close to 75 percent of the sources used in terms of being cited and consequently referenced, respectively. Thus the students in this English composition course relied most heavily on five types of sources in their references and used them with varying degrees of frequency as exemplified by the number of times these sources were cited within the body of their papers.

	TABLE 8 Most Frequently *Cited* Subfacet Combinations*			
Subfacet Combination	**Translation: Author Identity; Editorial Process; Publication Purpose**	**Frequency**	**Percent**	**Cumulative Percent**
3F4F5F	Academic Professional; Peer-reviewed; Higher Education	383	18.6	18.6
3E4E5B	Applied Professional; Editor and Editorial Staff; Commercial	312	15.2	33.8
3C4A5C	Corporate Author; Self-published; Nonprofit	134	6.5	40.3
3C4A5B	Corporate Author; Self-published; Commercial	117	5.7	45.9
3C4E5B	Corporate Author; Editor and Editorial Staff; Commercial	116	5.6	51.6
3B4E5B	Layman; Editor and Editorial Staff; Commercial	110	5.3	56.9
3C4A5E	Corporate Author; Self-published; Government	89	4.3	61.2
3F4E5F	Academic Professional; Editor and Editorial Staff; Higher Education	82	4.0	65.2
3F4E5B	Academic Professional; Editor and Editorial Staff; Commercial	65	3.2	68.4
3B4D5B	Layman; Moderated Submissions; Commercial	47	2.3	70.7
3C4E5C	Corporate Author; Editor and Editorial Staff; Nonprofit	45	2.2	72.9
3D4F5F	Professional-Amateur; Peer-reviewed; Higher Education	36	1.7	74.6
*(as found in papers' text), first three quartiles (n = 12)				

TABLE 9
Most Frequently *Referenced* Subfacet Combinations*

Subfacet Combination	Translation—Author Identity; Editorial Process; Publication Purpose	Frequency	Percent	Cumulative Percent
3E4E5B	Applied Professional; Editor and Editorial Staff; Commercial	111	16.0	16.0
3F4F5F	Academic Professional; Peer-reviewed; Higher Education	109	15.8	31.8
3C4A5C	Corporate Author; Self-published; Nonprofit	47	6.8	38.6
3C4A5B	Corporate Author; Self-published; Commercial	43	6.2	44.8
3C4A5E	Corporate author; Self-published; Government	37	5.3	50.1
3C4E5B	Corporate Author; Editor and Editorial Staff; Commercial	32	4.6	54.8
3F4E5F	Academic Professional; Editor and Editorial Staff; Higher Education	30	4.3	59.1
3B4E5B	Layman; Editor and Editorial Staff; Commercial	29	4.2	63.3
3F4E5B	Academic Professional; Editor and Editorial Staff; Commercial	23	3.3	66.6
3C4E5C	Corporate Author; Editor and Editorial Staff; Nonprofit	15	2.2	68.8
3E4E5C**	Applied Professional; Editor and Editorial Staff; Nonprofit	15	2.2	71.0
3B4D5B	Layman; Moderated Submissions; Commercial	11	1.6	72.5
3F4E5C**	Academic Professional; Editor and Editorial Staff; Nonprofit	11	1.6	74.1
3Z4Z5Z**	Source Unknown; Source Unknown; Source Unknown	11	1.6	75.7

*(as found in bibliographies), first three quartiles (n = 14)
**Subfacet combinations not appearing in table 8

Because 75 percent of all subfacet reference combinations were represented by only 14 of all subfacet combinations out of the 70 we coded, we tested whether any of these 14 subfacet combinations had any impact on certain quantitative variables (number of citations; cumulative GPA; student age; cumulative hours; and paper GPA grade, which was created to normalize each paper's grade). Using a nonparametric Kruskal-Wallis test due to the non-normally distributed data in these variables, only the *number of citations* variable showed a significant difference in the mean rank of the values when compared to the subfacet combinations (X^2= 175.837, p = 0.00). Two subfacet combinations, specifically 3E4E5B (Applied Professional; Editor and Editorial Staff; Commercial) (n = 111) and 3F4F5F (Academic Professional; Peer-reviewed; Higher Education) (n = 109) found in the papers' references stood out more than any other type of subfacet combination. We rejected the null hypothesis for only the *number of citations* variable. This means that students writing these papers relied on information resources written by an applied professional published under editorial supervision for commercial purposes and resources written by an academic published via peer review for educational/academic purposes significantly more than any other resource among the top 14 subfacet combinations.

Discussion

This pilot study provides a statistically reliable portrait of the authorities present in student sources in English 102, a multisection composition course within the university's first-year writing program that helped us gain insight into student information habits and the types of authorities they treat as authoritative and reliable. This study also clarifies that student responses to that landscape of sources are remarkably consistent. There were no significant patterns of source usage that could be statistically associated with better papers (as represented by the grade they received). Demographic characteristics such as age, gender, and race/ethnicity had no statistically demonstrable impact on source selection; likewise, a student's cumulative grade point average (GPA) in coursework prior to English 102 was not associated with any changes in the types of sources selected. The lack of relationship between source usage and either GPA or paper grade was a significant and surprising finding, as the research team had anticipated that more academically successful students would likely use scholarly sources with greater frequency than their less academically successful classmates.

One of our more significant realizations was the importance of identifying and understanding a source's subfacets in relationship to one another. The three facets (Author Identity, Editorial Process, and Publication Purpose) are interdependent, and they combine to provide necessary context for analysis. This interdependency was particularly obvious in corporately authored (subfacet 3C) and self-published (subfacet 4A) works. To say that the second most common kind of author type that appeared in student bibliographies was corporate (28.3 percent) tells us very little about student conceptions of authority. To say that corporately authored, self-published works by nonprofit organizations (subfacet combination 3C4A5C) at 6.8 percent narrowly edged out similar works by both commercial entities (subfacet combination 3C4A5B) at 6.2 percent, and government entities (subfacet combination 3C4A5E) at 5.3 percent provides a somewhat clearer picture of student source selection; moreover, it provides us with an easy and value-neutral means of describing subtly distinct types of material.

Use of Academic Materials

The results of the pilot offered some encouragement for information professionals interested in promoting the use of academic materials, particularly peer-reviewed sources. Although publications with a higher education purpose (subfacet 5F) were less prevalent than nonacademic sources overall, peer-reviewed journal articles by academic professionals (subfacet combination 3F4F5F) appeared in student bibliographies at comparable rates to commercial, editorially reviewed pieces by journalists and other applied professionals (subfacet combination 3E4E5B), comprising 15.8 percent and 16 percent of all sources, respectively. Furthermore, students cited peer-reviewed journal article sources at a disproportionately higher rate. These items accounted for 15.8 percent of bibliographic sources, but 18.6 percent of all paper citations.

Less encouraging for information professionals was the finding that works composed by academic professionals and editorially reviewed for higher education purposes (subfacet combination 3F4E5F), which largely represents books produced by academic presses, accounted for a modest portion of total bibliographic sources (4.3 percent) and paper citations (4.0 percent). As lessons on searching the book catalog were not a requirement of English 102 at the institution featured in this study, this may be a natural result of the curriculum. If this finding is confirmed in future studies, librarians may need to decide whether and how they wish to address this deficiency as educators. Incorporating catalog searching into all informa-

tion literacy sessions for students of English 102 may alter this trend. It also is possible that the study institution needs to alter its collection practices if it hopes to increase the use of academic books by this student population. However, it may be possible that other, nonbook formats might be better suited to the research needs of today's lower-level undergraduate students, especially as the offering of online courses and programs particularly continue to proliferate in the nation's universities.

Use of Nonacademic Materials

The larger takeaway from this pilot study was the presentation in the data of the breadth and variety of nonacademic material that students used as sources, even in the context of a research paper that they developed over several instructional weeks. The subfacet 5F (Higher Education), which is applied to all material produced for a higher educational purpose, accounted for only 27.5 percent of all references in student bibliographies—a number that includes materials outside the traditional peer-reviewed articles and books by academic publishers mentioned above. Furthermore, despite the wide variations between individual student papers in terms of number of sources used and the subject matter being addressed, only one of the sixty papers in our sample restricted itself solely to 5F (Higher Education) sources. Engaging with the material classified as nonacademic by the taxonomy yielded some valuable insights into the information landscape as it is used by students.

Trends in Editorial Review

Another insight made possible by the use of the taxonomy was the need for information literacy professionals to engage more seriously with the concept of editorial review. This study demonstrated that, when students reach for nonscholarly sources, the sources they selected were most likely to have undergone some kind of editorial review (47 percent of sources were coded as subfacet 4E: Editor and Editorial Staff); and, of those editorially reviewed sources, the majority were produced for commercial publications (63.1 percent of 4E sources were also designated as 5B: Commercial). While academia has clearly defined standards for what constitutes peer review, what constitutes editorial review within and beyond academia is indistinct.

The team also was surprised by the broad range of for-profit publications that indicated they practice some form of editorial review. In a single student paper on the health effects of cell phone usage, editorially reviewed material for commercial purposes (subfacet combination 4E/5B) included such familiar and anticipated sources as CNN, *Men's Health* magazine, and *USA Today*, alongside less familiar and anticipated sources such as CNET, *NFC World*, and Phone Arena. In doing the work of coding papers like this one, the research team observed that the appellation of "editor" was expanding to encompass miscellaneous creative and managerial duties beyond the scope of what normally constitutes editorial work within traditional publishing models. At the same time, the review process itself was sometimes obfuscated, even by well-known and widely trusted traditional sources. This has implications for our practice as information literacy professionals: how can we meaningfully distinguish between different kinds of editorial review in evaluating a source or in teaching students to arrive at their own assessment? It also has implications for the taxonomy in that its use of a single category for "editorial review" may be too broad: this is a question the research team intends to address in future research.

The subfacet 4D (Moderated Submissions) also posed challenges to traditional approaches to assessing the authority of a source. This designation, which the taxonomy uses for any

publishing outlet that appears to subject material to review by someone other than the author, but that does not identify any editorial staff, covered a surprising range of material. Works as diverse as a master's thesis from Eastern Michigan University, a post appearing on tronc's ChicagoNow blogging platform, and a page written by a student intern for the nonprofit Florida Innocence Project all received that designation. This reinforced for us that, regardless of how our profession addresses the questions surrounding editorial review, students cannot treat the presence or absence of an editorial board as a clear delineation between high- and low-quality sources. The combination of subfacets, though, provided some useful insight into the character of a 4D level of review for a particular source, since, in the case of the three examples listed above, each one was composed by a different subfacet 3 author for a different subfacet 5 purpose.

Use of Lay Authors
The taxonomy revealed the prevalence of lay authors in areas of the information landscape that we had not anticipated. While we had expected that students would use material written by laypeople with no professional or academic qualifications, it was a surprise to see the wide variety of publications that offer platforms to lay authors (subfacet 3B). Nonprofit (15.5 percent of 3B sources) and higher education sources (12.7 percent of 3B sources) published material by lay authors with unexpected frequency, even though commercial outlets (subfacet 5B) published the majority of lay authors (59.2 percent of 3B sources). This suggests that one common approach to the classification of websites in the library literature—classifying them by top-level domain and encouraging students to use sites with .org and .edu URLs—is less likely to lead to well-credentialed authors than might be expected.[14]

Suggested Approaches and Limitations
When considered on the whole, the findings suggest that, to produce information-literate graduates, information literacy professionals at colleges and universities will need to design their instructional efforts to help students develop a model for analyzing and evaluating nonscholarly materials. The predominant approach, in which librarians and instructors focus on promoting and teaching the trustworthy academic resources made available by the library, is inadequate to prepare students for thoughtful engagement with the sources they actually use. Given our experience, the research team is confident that the taxonomy represents a sound foundation for the development of that model. The taxonomy is an economical method for expressing the key characteristics of a source, which can be used to establish the authority of that source. The simplicity of this three-faceted measure makes it very plausible that librarians and instructors could teach students to use the taxonomy themselves when selecting sources, and that the taxonomy could even form the basis of a simple rubric for the evaluation of student bibliographies.

We note here that the statements articulated above characterize our study population, but they cannot be applied with 100 percent statistical certainty to *all* students in English 102 at the university in question. All means reported in this paper indicate participant means, as no data could be collected from students who declined to participate in the study. A subsequent study will be necessary to assess whether study participants differ demographically from nonparticipants in any significant way. Furthermore, because this pilot study represents data collected at one research site, from students enrolled in parallel sections of a single course,

we caution against applying this paper's findings to other universities' English composition courses, or to student source use in other disciplinary settings, until additional research is able to assess the extent to which these findings can be generalized.

While the coding of the subfacets was confirmed by a level of intercoder reliability that met the minimum accepted standard for the social sciences, we also acknowledge that, because the strength of the taxonomy depends upon a high level of consistency and objectivity, further improvements in intercoder reliability are desirable. The research team is therefore considering modest modifications to the scope notes applied to each subfacet, particularly in the area of editorial review (subfacet 4E), as well as the creation of additional coder training materials, to improve rates of agreement in future analyses.

Future Research

The research team presents this research as a pilot study that explores what kinds of insights might be gained from applying the taxonomy of Leeder, Markey, and Yakel (as modified for this study's purposes) to a wide array of research papers and types of sources. We envision that the collection of data from composition courses at other institutions, as well as from courses in other disciplines, would be beneficial both in refining the use of the taxonomy and in gaining a better understanding of how consistently college students rely on particular types of sources in different contexts. Future research also might examine international data to determine whether the features of the information landscape that are evident in these findings are characteristic of the information literacy habits of American students or if those features are common to students in other countries as well.

Conclusion

In conducting this pilot study, the research team gained insights into the current information landscape and student constructions of source authority with implications for instruction and collection development. Through the process of researching and categorizing each student source according to the modified Leeder, Markey, and Yakel taxonomy, we delved into largely unfamiliar areas of the web and examined familiar resources with fresh eyes.

The taxonomy excelled in providing a method for analyzing the nonscholarly sources that constituted more than two-thirds of student references and citations, as well as revealing complexity in the ways traditional journalistic and scholarly sources establish their authority (such as in their employment of laymen as authors). The study clarified that student responses to today's complicated landscape of sources are remarkably consistent. There are no significant patterns of source usage that can be statistically associated with better papers (as represented by the grade the paper received). Students use the sources they use, in the same general proportions, regardless of their demographic or academic background. This challenges us to engage with our students in a more authentic fashion.

To the research team, Leeder, Markey, and Yakel's modified faceted taxonomy and its application represent a potential path toward articulating a level of information literacy impact (ILI) evident in the research papers of students. ILI could be a conceptual measure rather than a strict mathematical measure, similar to other impact variables found in the library and information sciences (as only one example, the Journal Impact Factor or JIF used by the Citation Index Database). If information literacy is envisioned "as extending the arc of learn-

ing throughout students' academic careers and as converging with other academic and social learning goals," then effective analytical tools are required to measure the impact of any information literacy initiatives undertaken by information professionals and educators in higher education institutions.[15] We believe that the modified faceted taxonomy is one of those tools.

The approach structured by the taxonomy is easily situated in the language and perspective of the *Framework for Information Literacy*. Authority is constructed; while that construction is affected by many factors, it is powerfully impacted by the interplay of the taxonomy's three key elements of author identity, editorial process, and the purpose of the publication. Authority is also contextual; it is contextual in that the nature of the information need may help to determine the type of authority required. The process of analyzing and classifying the sources in student bibliographies caused us to consider the contexts from which these online sources arose and the needs those sources anticipate. Engagement with the taxonomy creates opportunities for librarians and composition instructors to examine their own habits of source discovery and evaluation. Independent of what can be learned about our students and the sources they trust, the use of the taxonomy equips information literacy professionals to identify and understand better the kinds of questions that all information users need to ask themselves more rigorously than we are generally taught to do.

APPENDIX A. Interpretation Guide

Facet 3: Author Identity

Author identity is connected to the person who is named and is taking responsibility for the content. For example, a reviewer at WebMD who is listed in connection with an article's content is scored as the author, even if the original author is unknown.

3A: Unknown Authorship

Definition:	The source indicates that the piece was anonymously written.
Scope Notes:	Known pseudonyms at the time the student accessed the content, which cannot be connected to a named individual, meet this standard.

3B: Layman

Definition:	A person without demonstrated expertise in the area being written about.

3C: Corporate Authorship

Definition:	No single author identified on a work issued by an organization.
Scope Notes:	Content attributed to a committee or other collective group is treated as 3C, even if the names of the members of that group are supplied. | An episode of a TV/radio series will be credited to the network/content provider that aired the episode.

3D: Professional-Amateur

Definition:	A person with a graduate degree in another field, but demonstrating interest, dedication, or experience in the area being written about.
Clarifying Example:	Someone with a PhD in philosophy who regularly writes about genetically modified foods on the web.

3E: Applied Professional

Definition:	A person with relevant experience, training, or credentials relevant to the area being written about.
Clarifying Example:	A journalist with a journalism degree OR substantive professional experience.
Scope Notes:	"Substantive professional experience" for a journalist is defined as service for at least one year in a writing staff position for a newspaper or magazine that publishes investigative reporting under the direction of an editorial board. | Graduate degree candidates in fields relevant to the area, who have not yet earned their degrees, are in this category.

3F: Academic Professional

Definition:	A person with a master's or doctoral degree in the area being written about, which they held at the time the content was published.
Clarifying Example:	PhD, MD, JD, MFA, MBA, and the like. For multiauthored pieces, at least one person must hold this degree.
Scope Notes:	If the author has the rank of a professor (Assistant, Associate, or Full), they meet this standard, with the exception of Visiting Professors, who often do not hold the same kind of academic credential.

3Z: Unknown

Definition: It was not possible to identify the source based on the available information in the citation.

Facet 4: Editorial Process

References to "the author" in Facet 4's notes refer to the person or entity scored by Facet 3.

4A: Self-published

Definition: Material made public directly by the author.

Clarifying Example: A personal blog.

4B: Vanity Press

Definition: Material the author paid to publish, generally as self-promotion.

4C: Collaborative Editing

Definition: Material that is reviewed or edited by multiple (possibly anonymous) collaborators.

Clarifying Example: A Wikipedia article.

4D: Moderated Submissions

Definition: Contributed content that has been accepted or approved by someone other than the author but that has not undergone editorial review.

Scope Notes: Content qualifies as 4D if it does not appear to be self-published, it does not appear that the author paid for publication (as vanity presses do), and it was not apparently produced in a collaborative editing structure (like a wiki or a group-authored blog), and that is credited to one or more named authors but is not reviewed by a professional editor or editorial staff as described in 4E. | When the content is by a corporate author on their own website, 4A is the appropriate category, since they are self-publishing that material. Corporately authored pieces, which otherwise would qualify as 4A, are qualified as 4E if a named editor is associated with the content.

4E: Editor and Editorial Staff

Definition: Professionally reviewed and approved by an editor and/or editorial staff.

Scope Notes: When the content is by a corporate author on their own website, 4A is the appropriate category, since they are self-publishing that material. Corporately authored pieces, which otherwise would qualify as 4A, are qualified as 4E if a named editor is associated with the content. | Editorial/opinion pieces from outlets with an editorial team qualify as 4E (such as CNN's opinion blog, The Chart). This will be true unless explicit information is available indicating that editors have no authority over a given section. | An editor-in-chief without any other named editorial staff is treated as an honorific title and does not qualify as 4E: other titles, such as "content editor," "features editor," and the like, do qualify. | A piece from an edited volume, if the editors are explicitly

named, will qualify as 4E. | Science advisory boards (common to many science and health sites online) are not normally considered editorial staff, as there is no way to know what capacity those advisors work in, or how analogous it is to editorial work. WebMD is an unusual exception—they explicitly describe their editorial policy, including the role of their medical experts, and they also list them as "Medical Editors." Sites that provide that kind of clarifying detail qualify as 4E.

4F: Peer-reviewed

Definition: Professionally reviewed and approved by peer scholars.

4Z: Unknown

Definition: It was not possible to identify the source based on the available information in the citation.

Facet 5: Publication Purpose

5A: Personal

Clarifying Example: A comment by a Goodreads user.

Scope Notes: Material produced for a political party or political organization in the United States that is not a 501(c) registered entity qualifies as 5A, as it does not meet the standard for 5C, but cannot be said to have a "commercial" purpose.

5B: Commercial

Clarifying Example: BBC Worldwide, as the commercial arm of the BBC, qualifies for this designation.

5C: Nonprofit

Scope Notes: If the organization is based in the United States, it must have 501(c) organizational status to qualify.

5D: K–12 Education

5E: Government

Clarifying Example: BBC's traditional news service, which is owned and operated by the government, qualifies for this designation.

5F: Higher Education

Scope Notes: All peer-reviewed journals qualify as 5F. | Publishers who cater primarily and explicitly to the higher education market (such as Routledge) qualify as 5F, regardless of whether the publisher is for-profit or nonprofit.

5Z: Unknown

Definition: It was not possible to identify the source based on the available information in the citation.

Notes

    1. Association of College & Research Libraries (ACRL), *Framework for Information Literacy for Higher Education*, available online at www.ala.org/acrl/standards/ilframework [accessed 11 January 2016].
    2. Ibid.

3. Chris Leeder, Karen Markey, and Elizabeth Yakel, "A Faceted Taxonomy for Rating Student Bibliographies in an Online Information Literacy Game," *College & Research Libraries* 73, no. 2 (2012): 115–33, doi:10.5860/crl-223.

4. Philip M. Davis, "Effect of the Web on Undergraduate Citation Behavior: Guiding Student Scholarship in a Networked Age," *portal: Libraries and the Academy* 3, no. 1 (2003): 41–51, doi:10.1353/pla.2003.0005; Jake Carlson, "An Examination of Undergraduate Student Citation Behavior," *Journal of Academic Librarianship* 32, no. 1 (2006): 14–22, doi:10.1016/j.acalib.2005.10.001; Stacey Knight-Davis and Jan Sung, "Analysis of Citations in Undergraduate Papers," *College & Research Libraries* 69, no. 5 (2008): 447–58, doi:10.5860/crl.69.5.447; David Mill, "Undergraduate Information Resource Choices," *College & Research Libraries* 69, no. 4 (2008): 342–55, doi:10.5860/crl.69.4.342; Yvonne Mery, Jill Newby, and Ke Peng, "Performance-Based Assessment in an Online Course: Comparing Different Types of Information Literacy Instruction," *portal: Libraries and the Academy* 12, no. 3 (2012): 283–98, doi:10.1353/pla.2012.0029; Kristina Howard, Thomas Nicholas, Tish Hayes, and Christopher W. Appelt, "Evaluating One-Shot Library Sessions: Impact on the Quality and Diversity of Student Source Use," *Community & Junior College Libraries* 20, no. 1 (2014): 27–38, doi:10.1080/02763915.2014.1009749; Carrie Ludovico and Carol Wittig, "A Universe of Information, One Citation at a Time: How Students Engage with Scholarly Sources," *Journal of Library & Information Services in Distance Learning* 9, no. 1/2 (2015): 30–39, doi:10.1080/1533290X.2014.946343.

5. Anne Middleton, "An Attempt to Quantify the Quality of Student Bibliographies," *Performance Measurement and Metrics* 6, no. 1 (2005): 7–18, doi:10.1108/14678040510588553; Caroline Cason, Barratt, Kristin Nielsen, Christy Desmet, and Ron Balthazor, "Collaboration Is Key: Librarians and Composition Instructors Analyze Student Research and Writing," *portal: Libraries and the Academy* 9, no. 1 (2009): 37–56; Rachel Cooke and Danielle Rosenthal, "Students Use More Books after Library Instruction: An Analysis of Undergraduate Paper Citations," *College & Research Libraries* 72, no. 4 (2011): 332–43, doi:10.5860/crl-90; Michelle Hendley, "Citation Behavior of Undergraduate Students: A Study of History, Political Science, and Sociology Papers," *Behavioral & Social Sciences Librarian* 31, no. 2 (2012): 96–111, doi:10.1080/01639269.2012.679884.

6. Deborah J. Grimes and Carl H. Boening, "Worries with the Web: A Look at Student Use of Web Resources," *College & Research Libraries* 62, no. 1 (2001): 11–23, doi:10.5860/crl.62.1.11; Sarah Blakeslee, "The CRAAP Test," *LOEX Quarterly* 31, no. 3 (2004): 6–7; Marc Meola, "Chucking the Checklist: A Contextual Approach to Teaching Undergraduates Web-Site Evaluation," *portal: Libraries and the Academy* 4, no. 3 (2004): 331–44, doi:10.1353/pla.2004.0055; Paul Stapleton and Rena Helms-Park, "Evaluating Web Sources in an EAP Course: Introducing a Multi-Trait Instrument for Feedback and Assessment," *English for Specific Purposes* 25, no. 4 (2006): 438–55, doi:10.1016/j.esp.2005.11.001; Rena Helms-Park, Pavlina Radia, and Paul Stapleton, "A Preliminary Assessment of Google Scholar as a Source of EAP Students' Research Materials," *Internet and Higher Education* 10, no. 1 (2007): 65–76, doi:10.1016/j.iheduc.2006.10.002; Davida Scharf, Norbert Elliot, Heather A. Huey, Vladimir Briller, and Kamal Joshi, "Direct Assessment of Information Literacy Using Writing Portfolios," *Journal of Academic Librarianship* 33, no. 4 (2007): 462–77, doi:10.1016/j.acalib.2007.03.005; Randall McClure and Kellian Donna Clink, "How Do You Know That? An Investigation of Student Research Practices in the Digital Age," *portal: Libraries and the Academy* 9, no. 1 (2009): 115–32, doi:10.1353/pla.0.0033; Thomas L. Reinsfelder, "Citation Analysis as a Tool to Measure the Impact of Individual Research Consultations," *College & Research Libraries* 73, no. 3 (2012): 263–77, doi:10.5860/crl-261; Alex P. Watson, "Still a Mixed Bag," *Reference Services Review* 40, no. 1 (2012): 125–37, doi:10.1108/00907321211203685; Helen Georgas, "Google vs. the Library (Part III): Assessing the Quality of Sources Found by Undergraduates," *portal: Libraries and the Academy* 15, no. 1 (2015): 133–61, doi:10.1353/pla.2015.0012; Holly Luetkenhaus, Steve Borrelli, and Corey Johnson, "First Year Course Programmatic Assessment," *Reference & User Services Quarterly* 55, no. 1 (2015): 49–60; Catherine Lantz, Glenda Maria Insua, Annie R. Armstrong, and Annie Pho, "Student Bibliographies: Charting Research Skills Over Time," *Reference Services Review* 44, no. 3 (2016): 253–65, doi:10.1108/RSR-12-2015-0053.

7. Andrew M. Robinson and Karen Schlegl, "Student Bibliographies Improve When Professors Provide Enforceable Guidelines for Citations," *portal: Libraries and the Academy* 4, no. 2 (2004): 275–90, doi:10.1353/pla.2004.0035; Davis, "Effect of the Web on Undergraduate Citation Behavior."

8. Stephanie Rosenblatt, "They Can Find It But They Don't Know What to Do with It: Describing the Use of Scholarly Literature by Undergraduate Students," *Journal of Information Literacy* 4, no. 2 (2010): 50–61, doi:10.11645/4.2.1486.

9. Leeder, Markey, and Yakel, "A Faceted Taxonomy for Rating Student Bibliographies."

10. Georgas, "Google vs. the Library (Part III)," 139.

11. Sarah P.C. Dahlen and Kathlene Hanson, "Preference vs. Authority: A Comparison of Student Searching in a Subject-Specific Indexing and Abstracting Database and a Customized Discovery Layer," *College & Research Libraries* 78, no. 7 (2017): 878–97, doi:10.5860/crl.78.7.878.

12. ACRL, *Framework for Information Literacy for Higher Education.*

    13. Andrew F. Hayes and Klaus Krippendorff, "Answering the Call for a Standard Reliability Measure for Coding Data," *Communication Methods and Measures* 1, no. 1 (2007): 77–89, doi:10.1080/19312450709336664.

    14. This approach appears in numerous published studies, including Middleton, "An Attempt to Quantify the Quality of Student Bibliographies"; Barratt et al., "Collaboration Is Key"; Cooke and Rosenthal, "Students Use More Books after Library Instruction"; and Hendley, "Citation Behavior of Undergraduate Students."

    15. ACRL, *Framework for Information Literacy for Higher Education.*