Citation Patterns of Engineering, Statistics, and Computer Science Researchers: An Internal and External Citation Analysis across Multiple Engineering Subfields

Madeline Kelly

This study takes a multidimensional approach to citation analysis, examining citations in multiple subfields of engineering, from both scholarly journals and doctoral dissertations. The three major goals of the study are to determine whether there are differences between citations drawn from dissertations and those drawn from journal articles; to test a methodology incorporating both internal and external citation sources; and to explore the citation habits of researchers in science, technology, engineering, and mathematics (STEM) subfields. The results reveal variations in how STEM subfields conduct research in career and academic settings and are more nuanced than internal or external citation data alone can provide. The results have practical collection development implications.

Library materials budgets are complex, allocating funds based on any number of factors: variables like faculty size, full-time enrollment, the average cost of journals, and expectation of program growth. One component often used to shape budget allocations is the serial/monograph ratio, or the proportion of serials, monographs, and other materials used by researchers in a given subject or discipline. Librarians have been looking at this ratio (if not calling it by name) since the 1920s and have written dozens of "characteristics of the research literature" articles.[1] Gross and Gross debuted the practice in their article, "College Libraries and Chemical Education."[2] Sixty years later, Devin and Kellogg compiled the findings of many of these earlier studies and proposed a consistent methodology for establishing the serial/monograph ratio and incorporating it into the budgeting process.[3] Over time, an increasing number of researchers have published studies building on this work, and libraries continue to use the results as a way to parcel out funding.

Nevertheless, few existing studies have used a methodology nuanced enough to be both locally applicable and widely relevant. Most are either too specific or too broad to be generalizable, resulting in serial/monograph ratios (and other related data) that

Madeline Kelly is Head of Collection Development in the University Libraries at George Mason University; e-mail: mkelly25@gmu.edu.

doi:10.5860/crl.76.7.859

crl14-600

cannot be easily used. The purpose of this study, which will be expanded upon in the following sections, is to compare various approaches to determining the serial/monograph ratio, to blend the strategies used by previous researchers into a single methodology that addresses both local and global need, and to yield usable data on the citation habits of science, technology, engineering, and math (STEM) researchers. Moving forward, the resulting model should help libraries determine a sound serial/monograph ratio and allocate their resources accordingly.

Literature Review

Historically, the serial/monograph ratio has been determined through citation analysis of either the broader scholarly literature or local research in the form of theses and dissertations. Kuruppu and Moore describe these two approaches as "global" and "local"; in this study they are called "external "and "internal" citation analysis.[4] A quick scan of the literature on citation analysis yields dozens of articles documenting studies in a variety of fields and contexts; the purpose of this literature review is not to be exhaustive, but merely to illustrate the range of approaches taken with both external and internal citation analysis projects.

External citation analyses have varied in both the source materials they use and the subject scope that they take. As for source, most studies have opted to analyze citations from top journals in the field.[5] Others have sampled citations from monographs, particularly in the humanities, where monographs play a large role in scholarly work.[6] Some studies have even chosen their source material on an individual basis, using works by specific, highly specialized researchers.[7] Whatever the source, these researchers and their published works are not specific to any particular institution. Rather, they represent the global pool of research, hence the term "external citation analysis."

The scope of these studies range from in-depth analysis of single subjects to generalized subject analysis to interdisciplinary comparisons. Single-subject studies have covered business computing, composition studies, linguistics, tourism, civil engineering, athletic training, international relations, and more.[8] Other studies have sought to explore multiple facets of a broad subject: Waller explored three subdivisions of economics, while the Medical Library Association coordinated a series of studies to map the literature of nursing across topics like public health, maternal/gynecologic nursing, and medical-surgical nursing.[9] Finally, some studies have taken a more general approach, looking at topics as broad as "the humanities," comparing general sciences to the humanities, or looking at variations among core humanities fields.[10] External citation analysis can follow any of these models, alone or in combination, to arrive at a cross-section of the desired group of researchers.

Internal citation analysis (analysis of locally produced research) varies in similar ways: there is leeway in both the source of the citations and the scope of the analysis. In this case, the most straightforward sources are theses and dissertations.[11] Some studies opt specifically for doctoral research while others favor master's-level work.[12] Other options for source material include the publications of local researchers and the work of undergraduate students.[13] Some of the most desirable source materials, however, are the publications of faculty, since (unlike students) faculty are present at universities for long stretches of time and represent a richer, more stable data pool. Finally, in rare instances, work can be gathered from sources at multiple institutions (such as undergraduates from four universities across the United States) for a means of comparison.[14] The unifying factor for all these source materials is that the researchers and their works are selected because of their institutional (or "internal") affiliation.

The scope of internal citation studies follows the same pattern as external analyses. Single subjects studied include psychology, biology, history, and education.[15] A few studies have compared academic programs or even broken programs into their component specializations.[16] Others have looked across the curriculum, comparing broad subjects like science and social science or English, education, and history.[17] The needs of the institution in question usually dictate the scope of the research.

For the purpose of this study, which explores STEM research (particularly engineering), all citation analyses that could be found on engineering topics were examined carefully. Among the external studies, a handful have looked at engineering or engineering-related fields: Devin and Kellogg analyzed engineering and electrical engineering citations; Larivière, Archambault, and Gingras compared engineering to other science and humanities subjects; and Curtis compared his own analysis of civil engineering citations to previous studies on the topic.[18] Additionally, Holsapple et al. explored business computing citations, which have some relevance to computer engineering.[19] Internally, St. Clair and Magrill studied undergraduate-level engineering citations; Williams and Fletcher explored master's-level engineering; Kirkwood analyzed civil engineering; and Sinn looked at theses and dissertations in statistics.[20] All of these studies provide at least some insight into the research habits of assorted STEM fields, though the variations in scope, methodology, and time period make them difficult to compare.

Finally, it is worth noting that, in addition to yielding the serial/monograph ratio, citation analysis typically provides data on the age of materials cited and the distribution of journals within a given field. Some researchers have also analyzed patterns in publishers or language of materials or have even parsed out the Library of Congress classes of the materials to better understand inter- or multidisciplinary subjects.[21] A citation analysis study can, therefore, inform not only budget decisions but collection development and collection assessment decisions as well.

Unfortunately, while citation analysis is a useful way to determine the serial/monograph ratio (and gather other relevant information), very few of the studies described above are more than tangentially applicable to any particular institution's engineering budget lines. The structure of university departments and library budgets are simply too varied. With a few exceptions, the available studies are either too general or too idiosyncratic to be applied. Larivière, Archambault, and Gingras, for instance, studied engineering citations—but only as a single, broad subject.[22] No distinctions were made among the various subfields of engineering, which limits the data's usability in a budgeting context. Thompson deals only with "the humanities"; Smith breaks things down only marginally further, into arts and humanities, education, science, and social science.[23] Even St. Clair and Magrill, with their extensive list of subjects surveyed, rely on undergraduate papers for their citations, rendering the data too soft to be useful, since undergraduates are unlikely to search exhaustively for the most appropriate resources, relying instead of materials that are easy to find.[24] These highly generalized studies are adequate for assessing overall trends, but insufficient for specific collection development or budgeting activities.

At the other end of the spectrum, overly specific studies focus on specialized subjects like business computing, composition, structural biology, or medical-surgical nursing.[25] As with the general studies, these detailed analyses have their place. Unfortunately, they go beyond the level of granularity necessary for budgeting in a large library setting. Indeed, it would be impractical to package this information meaningfully. Even the few engineering-specific studies are too limited, covering civil engineering but not, for instance, environmental engineering.[26] In short, few of these studies can be compared to one another, and the data are unlikely to correspond one-to-one to a library's subject fund lines. Even those that might be applicable are insufficient, since they fail to account for the citation habits of an institution's own research population.

Approximations are better than no data, but they are not ideal. There is no study that can provide libraries with a ready-to-use serial/monograph ratio.

In light of this inadequacy of the literature, the goals of this study—which were outlined roughly, above—are threefold: First, to ascertain whether internal and external citation analysis yield differing results, even within a single, specific subject area. If they vary significantly, then it can be supposed that neither is adequate on its own as a basis for collection development decisions; libraries must use both or risk having an unbalanced collection. Second, this study will test a methodology that incorporates both internal and external citations for a more nuanced result, mitigating the weaknesses of each approach. The institution-specific aspect of the internal citations will balance the overly general aspect of the external, and vice versa. If the methodology is successful, results will reflect local need without bowing to it completely. Other researchers will then be able to replicate the approach, sharing what data can be shared and gathering internal data as needed. Third, this study will examine the citation habits of STEM researchers, generating serial/monograph ratios and other data that can be used directly by libraries with engineering budget lines. Unlike many studies that present internal- or external-based case studies, this study uses the author's home institution, George Mason University (Mason), as an example to address the wider fundamental problem of using citation data from just one kind of source. In short, this study seeks to provide a more nuanced look at engineering citations than previous studies—one that applies to Mason specifically, but that can easily be mapped onto other institutions.

Methodology

This study was conducted at George Mason University (Mason), a medium-sized public university in Northern Virginia with an enrollment of just over 30,000 graduate and undergraduate students. The University Libraries—comprising a central library, three distributed libraries, and a law library—hold more than 1.4 million volumes, 3.3 million items on microform, and 11,000 print periodical subscriptions.[27] Collection development responsibilities are shared among more than 20 subject selectors and are coordinated by a centralized collection development department. Because Mason, like most universities, is such a complex parent organization, many factors go into the development of the library collections budget, just as countless factors go into the development of the collection itself. This organizational structure makes Mason a good setting in which to explore the serial/monograph ratio and the related collection development implications of citation analysis.

The first step in conducting this citation analysis, designed to ensure applicability to Mason, was to identify programs in the university's Volgenau School of Engineering that are represented by individual library budget lines. These areas were: bioengineering; computer science; applied information technology; electrical engineering; civil and environmental engineering; systems engineering and operations research; and statistics. (It can be argued that statistics does not belong among engineering fields; however, for the sake of thoroughness, all Volgenau programs with specific funding were included for study.) Second, engineering programs with more than one discrete subject area (for instance, "civil and environmental engineering") were subdivided into their component parts (like "civil engineering" and "environmental engineering"). The purpose of this step was to create the most basic possible engineering categories out of Volgenau's curriculum. This conceptual granularity opens the results up to other institutions, even if their specific programs are structured differently from those at Mason. For example, an institution with a robust civil engineering program but no focus on environmental engineering would be able to use the civil engineering data without interference from any environmental engineering component. In the end, eight engineering subfields emerged that applied equally to Mason and to the wider community: bioengineering, computer science, civil engineering, electrical engineering, environmental engineering, operations research, statistics, and systems engineering.

Within this framework, the study analyzed two kinds of citations: citations from the broader scholarly literature and citations from Mason doctoral dissertations. (While it would have been preferable to use faculty publications as the internal citation pool for this study, dissertations were easier to sample and thus more feasible given the time constraints of the project.) This two-pronged approach was intended to provide Mason with a more finely tuned serial/monograph ratio that included global as well as local research trends. Data were gathered as follows: First, eight journals were selected to represent the subfields in question. This first set of journals was identified using Thomson Reuters Web of Knowledge, ISI Journal Citation Index, and SciMago Journals and Country Rank. Selections were based on impact factor, five-year impact factor, cites per document, and the Article Influence Score. Mason's engineering librarian was also consulted to verify that journals were appropriate in scope and level. The resulting eight titles were deemed to be peer-reviewed exemplars of each particular subfield.

The first round of journals included: *Annals of Biomedical Engineering* (bioengineering); *Journal of Civil Engineering and Management* (civil engineering); *Computer Journal* (computer science); *Proceedings of the IEEE* (electrical engineering); *Journal of Environmental Engineering* (environmental engineering); *Operations Research* (operations research); *Annals of Statistics* (statistics); and *Systems Engineering* (systems engineering). Articles were sampled from the 2012 volume of each of these eight journals to provide relatively current results. The number of articles chosen from each journal was based on the average number of citations per article, with a goal of at least 1,000 total citations per subfield. (This pool of citations, with the addition of 500 more at a later stage—see below—was sufficient to provide results with 95 percent confidence and a margin of error of approximately 2.5. While more exact results would have been possible given more citations, there were practical limitations against gathering a larger pool.) The bibliography of each article sampled was then copied into a spreadsheet, along with the journal title, issue number, and source article publication data. Each subfield was allotted its own spreadsheet to facilitate processing and analysis.

Citations were then manually tagged with a format, year of publication, journal title, monograph title, and publisher. Formats included book, book chapter, conference proceeding, dissertation, journal, manuscript (for unpublished personal articles), patent, personal (for items like e-mail communications and interviews), presentation (for presentations not gleaned from conference proceedings), software, standard, technical report (which included technical reports, short government documents, commercial reports, working papers, and other gray literature), and website. All formats except "journal" and "personal" were labeled with publishers (when possible); only books and conference proceedings were labeled with a monograph title. All journal titles were recorded. This process was carried out both by the author and by an undergraduate student; student work was checked for mistakes and inconsistency using various "sort" functions in Excel.

Once the first pool of citations had been processed and corrected, a second, smaller group of citations was gathered from top journals identified during the first round of analysis. These secondary journals were chosen based on the number of authors citing them in the first pool of citations, as well as their scope relative to the subfield in question. Articles and citations were sampled from these secondary journals following the same procedures as the first group of journals. During this round of sampling, enough articles were selected to yield at least 500 citations per subfield (added to the first round of analysis, this created a cumulative pool of 1,500 citations per subfield).

The citations were then processed in the same way as the first batch; once complete, all citations were combined into a single master spreadsheet and corrected again to eliminate variant titles and misspellings. All data analysis was conducted using the full pool of 1,500 citations.

The second round of journals included: *Journal of Biomechanics* (bioengineering); *Engineering Structures* (civil engineering); *Communications of the ACM* (computer science); *Electronics Letters* (electrical engineering); *Environmental Science and Technology* (environmental engineering); *Management Science* (operations research); *Journal of the American Statistical Association* (statistics); and *IEEE Transactions on Systems, Man, and Cybernetics A* (systems engineering).

The final group of citations was sampled from Mason dissertations in the five engineering programs with established doctoral programs: Applied Information Technology (AIT), Computer Science (CS), Electrical Engineering (ECE), Systems Engineering and Operations Research (SEOR), and Statistics (STAT). Dissertations were retrieved from ProQuest Dissertations and Theses; the target was to retrieve 20 dissertations from each program (for a total of approximately 1,200 citations each). For CS, AIT, and ECE, there were sufficient dissertations from the years 2008–2012 to choose a random sample. Ultimately, 19, 22, and 18 dissertations were gathered for these subjects, respectively.

For SEOR, there were too few dissertations during the 2008–2012 timeframe to make an adequate sample, so the range was expanded to include 2005–2012. Finally, there were so few STAT dissertations available that two strategies were adopted to gather a large enough pool: First, the timeframe for STAT was expanded to include 2005–2012. Second, dissertations were drawn from a program other than STAT that includes pure statistics as a concentration: Computational Science and Informatics (CSI). Irrelevant CSI dissertations were avoided by selecting only those with the single subject heading “statistics” in ProQuest Dissertations and Theses. The sampling methods for SEOR

TABLE 1
Number of Citations Sampled

	Total Citations	Citations per Article
External (Journal) Citations		
Bioengineering	1,892	46.15
Civil Eng.	1,747	34.25
Computer Science	1,540	27.50
Electrical Eng.	1,689	23.46
Environmental Eng.	1,818	40.40
Operations Research	1,935	40.31
Statistics	1,724	33.80
Systems Eng.	1,804	41.00
Internal (Dissertation) Citations		
Computer Science (CS)	1,965	103.42
Electrical & Computer Eng. (ECE)	1,692	94.00
Sys. Eng. & Op. Research (SEOR)	1,736	86.80
Statistics (STAT)	1,242	62.10
Applied Info. Technology (AIT)	1,859	84.50

and STAT were not ideal, but they should still provide some subject-specific insight. SEOR and STAT included 20 dissertations each, with 8 of the STAT dissertations coming from CSI.

Citations were copied from the dissertations' bibliographies and processed as with the journal citations, above. At least 1,200 citations were collected for each subfield, allowing for results with 95 percent confidence and a margin of ±2.8. Additional categories for format were added to the results as needed to accommodate scholars' sources: interviews, personal (including personal correspondence and class lecture notes), and RFC (request for comment; there were enough RFCs to warrant a separate category). Ultimately, many of these miscellaneous formats were consolidated into the broad category, "other," for analysis.

Results

As anticipated, citations from the eight subfields that were studied exhibit distinct patterns in terms of format, currency, and journal preferences. There are also notable differences between the results drawn from external citations and those from internal citations, suggesting that external and internal citation analyses provide different perspectives even if conducted in identical fields.

Format of Cited Resources

Because this study was undertaken with the serial/monograph ratio in mind, format is the primary area of interest when exploring the results. Among the external citations, journals are the dominant format, ranging from 40 percent of citations (in computer science) to 94 percent (in bioengineering). Books and conference proceedings are the next most common formats. The percentage of books ranges from 3 percent (in bioengineering) to 25 percent (in systems engineering). In this case, systems engineering is an outlier; in most subfields the percentage of book citations falls considerably lower. Conference proceedings are widely dispersed, with a minimum of 2 percent (bioengineering) and a maximum of 39 percent (computer science). It is worth noting that computer science citations are as frequently conference proceedings as they are journals. Other formats make up a relatively small percentage of each subfield's external citations.

Among the internal citations, the formats are far more broadly distributed than in the external results. Journals remain a majority format, except in CS and AIT, where they are surpassed slightly by conference proceedings. Specifically, journals range from 27 percent of citations (AIT) to 61 percent (ECE) which, while substantial, is still much lower than among the external citations. Conferences are highly represented in many subjects, ranging from 6 percent of citations (STAT) to 43 percent (CS). In most subjects, the percentage falls near the middle, and is higher than among the external citations, where the median and mode for conferences were less than 6 percent. The percentage of books is comparable to the external citations, if slightly higher. Reliance on technical reports (and other gray literature) is also comparable between internal and external citations: with the exception of SEOR and AIT, engineering doctoral students are citing technical reports less than 6 percent of the time. Websites enjoy a slight upturn among the internal citations, though not consistently across all subfields. The results confirm the hypothesis that researchers in distinct fields rely on formats differently, as do researchers working in different settings.

Age of Cited Resources

In addition to revealing format preferences, this citation analysis yielded the ages of resources cited, which were as varied as the formats used. In the external literature,

TABLE 2
Format of Materials Cited

	Book/Book Chapter	Conference Proceeding	Journal	Website	Other
External (Journal) Citations					
Bioengineering	3%	2%	94%	<1%	<1%
Civil Engineering	12%	8%	69%	1%	10%
Computer Science	10%	39%	40%	5%	6%
Electrical Engineering	6%	27%	58%	5%	4%
Environmental Eng.	9%	3%	78%	2%	8%
Operations Research	15%	3%	78%	1%	3%
Statistics	14%	3%	75%	1%	7%
Systems Engineering	25%	13%	48%	4%	10%
Internal (Dissertation) Citations					
Computer Sci. (CS)	12%	43%	30%	10%	5%
Elec. & Comp. Eng. (ECE)	12%	20%	61%	2%	5%
Sys. Eng. & Op. Res. (SEOR)	22%	16%	30%	8%	26%
Statistics (STAT)	25%	6%	59%	4%	6%
Applied Info. Tech. (AIT)	16%	30%	27%	13%	14%

while all of the subfields cited resources more than 50 years old, the vast majority of resources cited (>98%) were published within the last 50 years. For all subfields except operations research, more than two thirds of cited materials were published in the last 15 years. Operations research favors the widest age range, with almost a fifth of resources being older than 25 years. Statistics and civil engineering are similar. The subfields that cite the most recent materials are computer science, bioengineering, and electrical engineering. In these subfields, materials published in the last 15 years account for 80 percent or more of citations. Electrical engineering relies on the newest resources, with almost half of citations pointing to materials published in the last 5 years.

As with the external citations, most internal citations (>96%) point to materials published within 50 years of the author's writing the dissertation. With the exception of STAT, approximately 80 percent of citations point to materials published within 25 years of the dissertation, and 75 percent or more are from the last 15 years. Computer science (CS) and AIT favor the newest resources. In AIT, 25 percent of citations are from the last two years, and 45.5 percent are from the last five years. Similarly, it is worth noting that for CS and SEOR, internal results show a much stronger preference for materials published within the last five years. In both subfields, 44 percent of citations are from the last five years; their external counterparts—computer science, operations research, and systems engineering—show 35 percent, 19 percent, and 30 percent, respectively. This contrasts with STAT, which varies far less between the internal and external results.

STAT favors the broadest range of publication dates, with only 10 percent from the two years prior to the publication of each dissertation and one fifth of materials older than 20 years. STAT also has the highest percentage of materials more than 50 years old. These results are similar to those found using external citations, although the internal citations for STAT are more evenly distributed along the age range than in the external citations. (It is worth noting that STAT is an outlier, but not surprisingly so, given that it relates more closely to mathematics than to engineering. A citation analysis of mathematics resources would undoubtedly put statistics in a different context.)

TABLE 3
Publication Date of Materials in Relation to Date Cited

	Within 2 Years	Within 5 Years	Within 10 Years	Within 15 Years	Within 25 Years	Within 50 Years	Within 75 Years	Within 100 Years
External (Journal) Citations								
Bioengineering	9%	32%	64%	80%	93%	99%	99.8%	99.9%
Civil Engineering	10%	34%	60%	74%	88%	99%	99.9%	99.9%
Computer Science	10%	35%	67%	82%	92%	99%	99.4%	99.7%
Electrical Engineering	13%	47%	73%	84%	93%	99%	99.6%	99.6%
Environmental Engineering	5%	27%	56%	75%	90%	99%	99.7%	99.8%
Operations Research	4%	19%	45%	64%	83%	98%	99.8%	99.8%
Statistics	10%	30%	54%	70%	88%	98%	99.9%	100.0%
Systems Engineering	5%	30%	59%	77%	90%	98%	99.0%	99.5%
Internal (Dissertation) Citations								
Computer Science (CS)	20%	44%	72%	85%	94%	99%	99.9%	99.9%
Electrical & Computer Engineering (ECE)	17%	38%	64%	76%	90%	99%	99.8%	99.9%
Systems Engineering & Operations Research (SEOR)	21%	44%	67%	81%	91%	99%	99.8%	99.9%
Statistics (STAT)	11%	27%	49%	62%	80%	96%	99.0%	99.8%
Applied Information Technology (AIT)	25%	46%	70%	84%	94%	99%	99.6%	99.6%

Journal Distribution among Cited Resources

A third factor illustrated by the results of this citation analysis is the number of journals cited in each field, and the frequency with which those journals are cited. Among the external results, the field with the greatest diversity of journal titles cited is computer science, with half of all journal citations indicating unique titles. Systems engineering is similarly diverse. The field with the smallest number of journals is statistics, where only 20 percent of journals cited are unique titles and 80 percent are duplicate references to the same set of core journals. Regardless of overall diversity, several subfields cite a small core of journals heavily. In statistics, 6 percent of journals are cited ten or more times. Other subfields with high concentrations of heavily cited journal titles include civil engineering, environmental engineering, and operations research.

Among the internal results, the field with the greatest journal diversity is AIT, with more than half of journal citations indicating a unique title. SEOR and CS follow. SEOR also has the highest percentage of journals that are cited in only one dissertation (86%), indicating a wide but relatively shallow pool of resources used by SEOR researchers. AIT follows, with 83.8 percent of journals being cited only once. The field with the smallest number of journals cited is ECE. As with the external results, STAT shows the highest concentration of heavily used journals, with 3 percent of journals cited by six authors or more and 0.7 percent cited by ten authors or more. This is notable, as the vast majority of journals—across all subjects—are cited in four dissertations or fewer. It also contrasts with the external results, where each subfield has at least one journal cited by ten authors or more.

TABLE 4
Frequency and Distribution of Journal Citations

	# of Unique Journals Cited	% of Journal Citations Representing Unique Titles	% Cited in 1 Article	% Cited in 2–5 Articles	% Cited in 6–9 Articles	% Cited in 10+ Articles
External (Journal) Citations						
Bioengineering	574	32%	71.6%	24.7%	2.8%	0.9%
Civil Engineering	334	28%	73.7%	21.9%	3.3%	1.1%
Computer Science	304	50%	78.0%	21.0%	0.3%	0.7%
Electrical Engineering	360	37%	79.7%	17.8%	1.7%	0.8%
Environmental Engineering	435	31%	72.2%	24.1%	2.3%	1.4%
Operations Research	382	25%	68.8%	24.9%	3.7%	2.6%
Statistics	264	20%	62.1%	28.4%	3.4%	6.1%
Systems Engineering	415	48%	80.0%	18.3%	1.2%	0.5%
Internal (Dissertation) Citations						
Computer Science (CS)	278	47%	78.8%	20.8%	0.4%	0.0%
Electrical & Computer Engineering (ECE)	228	22%	73.2%	25.0%	1.8%	0.0%
Systems Engineering & Operations Research (SEOR)	253	49%	86.2%	13.8%	0.0%	0.0%
Statistics (STAT)	303	41%	79.2%	17.8%	2.3%	0.7%
Applied Information Technology (AIT)	272	55%	83.8%	15.4%	0.8%	0.0%

Top Journals Cited

The final piece of information to be gleaned from the data is a list of core resources for each subfield. The most heavily cited journals from the external literature were identified for each subfield based on the number of authors citing each journal. While these titles are widely used, it is worth noting two things. First, the 16 journals used as source material for the external analysis appear within the top three ranks of each subfield; one cannot dismiss the possibility of some bias in their ranking. Second, some of the most heavily cited journals may still only receive middling support. In computer science, for instance, the top journal was cited in only 29 percent of the articles sampled. Other journals are more clearly influential: in statistics, for example, the *Annals of Statistics* was cited in 96 percent of articles. From an interdisciplinary standpoint, there were 12 journals that appeared in the literature of five or more of the subfields: *Science*, *Proceedings of the National Academy of Science, Nature, IEEE Transactions on Pattern Analysis and Machine Intelligence, Communications of the ACM, IEEE Transactions on Information Theory, New England Journal of Medicine, IEEE Transactions on Automation Control, Automatica, IEEE Transactions on Signal Processing, Lancet,* and *Journal of Theoretical Biology*. Because of their multidisciplinary appeal, these journals should be part of any large engineering collection.

In the internal literature, the same process was used to identify top journals. Unlike with the external citations, where bias in favor of the source journals may have skewed the results, the journals culled from dissertations probably represent a more neutral sample of top resources. These journals also represent research habits specific to Mason. All subfields yielded at least three top journals that coincided with the external list, boosting their status as core resources. From an interdisciplinary standpoint, there were only four journals that appeared in all five dissertation subfields (*IEEE Transactions on Pattern Analysis and Machine Intelligence, Annals of Mathematical Statistics, Machine Learning,* and *Operations Research*), while 18 were cited in four. Five of the titles from this interdisciplinary list overlap with the corresponding list for external citations: *Science, IEEE Transactions on Pattern Analysis and Machine Intelligence, Communications of the ACM, IEEE Transactions on Information Theory*, and *Automatica*. These lists of journals, while not the definitive core literature, represent a bare minimum for collection development efforts.

TABLE 5A **Top Bioengineering Journals,** **as Indicated by the Percentage of Authors Citing**
External
Journal of Biomechanics (54%)
Annals of Biomedical Engineering (44%)
Journal of Biomechanical Engineering (44%)
Circulation (27%)
Proceedings of the National Academy of Sciences (24%)
Biophysical Journal (22%)
Nature (22%)
American Journal of Physiology: Heart and Circulatory Physiology (20%)
IEEE Transactions on Biomedical Engineering (20%)
New England Journal of Medicine (20%)
Science (20%)

TABLE 5B
Top Civil Engineering Journals, as Indicated by the Percentage of Authors Citing

External
Engineering Structures (47%)
Journal of Civil Engineering and Management (43%)
Journal of Structural Engineering (29%)
Journal of Construction Engineering and Management (24%)
Journal of Constructional Steel Research (16%)
International Journal of Project Management (14%)
International Journal of Solics and Structures (14%)
Journal of Engineering Mechanics (14%)
Materials and Structures (14%)

TABLE 5C
Top Environmental Engineering Journals, as Indicated by the Percentage of Authors Citing

External
Water Research (60%)
Environmental Science and Technology (49%)
Journal of Environmental Engineering (40%)
Chemosphere (27%)
Water Science & Technology (27%)
Water Resources Research (22%)
Journal of Environmental Quality (20%)
Journal of Hazardous Materials (20%)
Science (20%)

TABLE 5D
Top Computer Science and Applied Information Technology Journals, as Indicated by the Percentage of Authors Citing

Computer Science		Applied Information Technology
External	**Internal**	**Internal**
Communications of the ACM (29%)	*Communications of the ACM (32%)*	*Communications of the ACM (36%)*
Science (18%)	IEEE Transactions on Pattern Analysis and Machine Intelligence (26%)	*IEEE Transactions on Knowledge and Data Engineering (27%)*
Computer Journal (16%)	*Journal of the ACM (26%)*	Journal of the American Statistical Association (23%)
ACM Transactions on Computer Systems (9%)	Machine Learning (26%)	*ACM Computer Surveys (18%)*

TABLE 5D
Top Computer Science and Applied Information Technology Journals, as Indicated by the Percentage of Authors Citing

Computer Science		**Applied Information Technology**
External	**Internal**	**Internal**
IEEE Transactions on Information Theory (9%)	*ACM Transactions on Computer Systems (21%)*	IEEE Computer (18%)
Journal of the ACM (9%)	ACM Transactions on Embedded Computer Systems (21%)	IEEE Transactions on Pattern Analysis and Machine Intelligence (18%)
Nature (9%)	IEEE Journal of Selected Areas in Communications (21%)	
ACM Computing Surveys (7%)	IEEE Transactions on Computers (21%)	
Communications (7%)	IEEE Transactions on Software Engineering (21%)	
Computer (7%)	*IEEE/ACM Transactions on Networking (21%)*	
IEEE Transactions on Image Processing (7%)		
IEEE Transactions on Knowledge and Data Engineering (7%)		
IEEE Transactions on Parallel and Distributed Systems (7%)		
IEEE/ACM Transactions on Networking (7%)		
Proceedings of the National Academy of Sciences (7%)		

*Italicized titles indicate overlap between external and internal results.

TABLE 5E
Top Electrical Engineering Journals, as Indicated by the Percentage of Authors Citing

External	**Internal**
Electronics Letters (21%)	*Proceedings of the IEEE (50%)*
IEEE Transactions on Antennas and Propagation (18%)	IEEE Transactions on Information Theory (39%)
Proceedings of the IEEE (15%)	IEEE Transactions on Signal Processing (33%)
Microwave and Opitcal Technology Letters (13%)	*Journal of Applied Physics (33%)*

TABLE 5E
Top Electrical Engineering Journals, as Indicated by the Percentage of Authors Citing

External	Internal
IEEE Transactions on Microwave Theory and Technology (10%)	*Applied Physics Letters (28%)*
IEEE Communications Magazine (8%)	IEEE Transactions on Communications (28%)
IEEE Transactions on Pattern Analysis and Machine Intelligence (8%)	IBM Journal of Research & Development (22%)
Journal of Applied Physics (8%)	IEEE Signal Processing Magazine (22%)
Proceedings of the National Academy of Sciences (8%)	IEEE Transactions on Acoustics, Speech, and Signal Processing (22%)
Applied Physics Letters (7%)	IEEE Transactions on Wireless Communication (22%)
Science (7%)	Solid-State Electronics (22%)
*Italicized titles indicate overlap between external and internal results.	

TABLE 5F
Top Systems Engineering and Operations Research Journals, as Indicated by the Percentage of Authors Citing

Systems Engineering	Systems Engineering & Operations Research	Operations Research
External	**Internal**	**External**
Systems Engineering (48%)	Air Traffic Control Quarterly (25%)	*Management Science (77%)*
IEEE Transactions on Systems, Man, and Cybernetics A (34%)	*Operations Research (25%)*	*Operations Research (67%)*
IEEE Transactions on Systems, Man, and Cybernetics C (20%)	IEEE Transactions on Automatic Control (20%)	*European Journal of Operations Research (38%)*
Harvard Business Review (14%)	*IEEE Transactions on Systems, Man, and Cybernetics (20%)*	American Economic Review (33%)
IEEE Software (14%)	Journal of the Operational Research Society (20%)	Manufacturing and Service Operations Management (31%)
IEEE Transactions on Systems, Man, and Cybernetics (14%)	*Management Science (20%)*	Mathematics of Operations Research (29%)
Management Science (14%)	*Systems Engineering (20%)*	Journal of Political Economy (25%)
Academy of Management Review (9%)	Transportation Science (20%)	Econometrica (23%)

TABLE 5F
Top Systems Engineering and Operations Research Journals, as Indicated by the Percentage of Authors Citing

Systems Engineering	Systems Engineering & Operations Research	Operations Research
External	**Internal**	**External**
Automatica (9%)	*European Journal of Operational Research (15%)*	Operations Research Letters (21%)
CrossTalk (9%)	IEEE Transactions on Software Engineering (15%)	Quarterly Journal of Economics (21%)
IEEE Systems Journal (9%)	Journal of Air Transport Management (15%)	
IEEE Transactions on Automation Science and Engineering (9%)	Journal of Guidance, Control, and Dynamics (15%)	
International Journal of Production Economics (9%)	Progress in Astronautics and Aeronautics (15%)	
Proceedings of the IEEE (9%)		
Science (9%)		

*Italicized titles indicate overlap between external and internal results.

TABLE 5G
Top Statistics Journals, as Indicated by the Percentage of Authors Citing

External	Internal
Annals of Statistics (96%)	*Journal of the American Statistical Association (60%)*
Journal of the American Statistical Association (63%)	*Annals of Statistics (50%)*
Biometrika (53%)	Biometrics (45%)
Journal of the Royal Statistical Society B (37%)	*Biometrika (40%)*
Bernoulli (31%)	Computational Statistics & Data Analysis (35%)
Journal of Multivariate Analysis (27%)	Statistics in Medicine (35%)
Scandinavian Journal of Statistics (27%)	Annals of Mathematical Statistics (30%)
Journal of Computational and Graphical Statistics (24%)	*Journal of Computational and Graphical Statistics (30%)*
Journal of Machine Learning Research (24%)	Journal of Statistical Planning and Inference (30%)
Journal of the Royal Statistical Society (24%)	*Journal of the Royal Statistical Society B (25%)*

*Italicized titles indicate overlap between external and internal results.

TABLE 6
Most Interdisciplinary Journals Based on the Number of Sub-Fields Citing

Title	# of Sub-Fields Citing	% of Authors Citing
External (Journal) Citations (n=8)		
Science	7	10.0%
Proceedings of the National Academy of Science	7	9.0%
Nature	6	7.5%
IEEE Transactions on Pattern Analysis and Machine Intelligence	6	4.0%
Communications of the ACM	5	6.0%
IEEE Transactions on Information Theory	5	5.0%
New England Journal of Medicine	5	3.5%
IEEE Transactions on Automation Control	5	3.0%
Automatica	5	2.5%
IEEE Transactions on Signal Processing	5	2.0%
Lancet	5	2.0%
Journal of Theoretical Biology	5	1.5%
Internal (Dissertation) Citations (n=5)		
IEEE Transactions on Pattern Analysis and Machine Intelligence	5	13.0%
Annals of Mathematical Statistics	5	12.0%
Machine Learning	5	12.0%
Operations Research	5	10.0%
Journal of the American Statistical Association	4	19.0%
Communications of the ACM	4	18.0%
IEEE Transactions on Information Theory	4	13.0%
Biometrika	4	11.0%
IEEE Transactions on Software Engineering	4	11.0%
IEEE Journal of Selected Areas in Communications	4	9.0%
IEEE Transactions on Systems, Man, and Cybernetics	4	9.0%
Journal of the Royal Statistical Society	4	9.0%
ACM Computer Surveys	4	8.0%
Science	4	8.0%
Automatica	4	7.0%

Discussion

As was predicted, there were wide variations among the subfields studied, as well as variations between internal and external analysis. Beyond emphasizing the need for both internal and external data, these variations have practical implications for collection development decisions.

Format of Cited Resources

The predominance of journals across nearly all subfields supports other citation studies—as well as anecdotal accounts—suggesting that science fields favor journals over monographs, especially in relation to the social sciences and humanities.[28] These patterns likely reflect the fast rate of change in engineering. Even STAT shows a high rate of journal usage, despite Sinn's assertion that statisticians use journals less frequently than other scientists do.[29] Actually, the opposite seems true: statistics falls at the upper end of journal use. In fact, statistics researchers appear to use both books and journals heavily, resulting in a relatively low serial/monograph ratio, but in high absolute usage. Engineers and their colleagues rely heavily on journals.

That said, these results show a remarkable range in the percentage of journal citations: 94 percent for bioengineering versus 40 percent for computer science. There are several ways to explore this. First, bioengineering is a relatively new field. Considering the age of most bioengineering materials (see table 3), it may simply be that there are still comparatively few bioengineering monographs. Similarly, it is also probably a fast-growing field, putting journals at the forefront of the ever-changing research. This idea of growth versus stability may also account for why some of the better-established fields (like statistics, or even systems engineering) cite a relatively high percentage of books versus journals. In these fields, there are more seminal tomes to rely on.

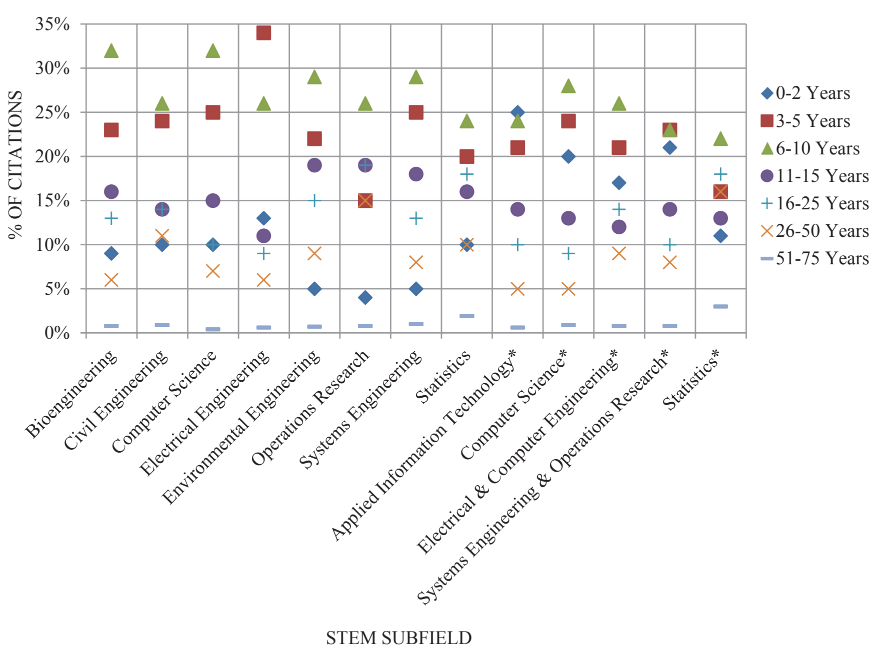
Computer science is an interesting exception to the pattern. For one of the fastest-paced branches of science and engineering, it relies relatively little on journals. The same is true of AIT and, to some extent, electrical engineering. Why? Computer science is constantly evolving; it also has a huge number of specialties, niches, and subcommunities. Both factors may contribute to the even split between journals and conferences, since both formats allow for a robust community to exchange information quickly in a topic-specific way. The high percentage of conferences may also have to do with the fact that many computer scientists are practitioners in the private sector and may prefer to trade information at conferences rather than publishing lengthy articles. They may also rely more heavily on non–peer-reviewed resources, which can be published more quickly and appear in sources other than journals. Whatever the precise cause, computer science exhibits its own pattern, depending more heavily on conference proceedings than other engineering fields.

Finally, it is important to note that journals are less cited in dissertations than in the wider literature—by a large margin. In fact, the formats cited in dissertations are far more evenly distributed than in the external citations; SEOR is an excellent example, with near-equal percentages of each major format. This dispersion reflects key features of the dissertation process, including the need to be exhaustive and the sheer length of a finished doctoral work. Further, the relative weakness of journal citations in dissertations may reflect the difference in publishing imperatives between PhD candidates and other researchers in the field.

For collection development, the patterns in format usage underscore the need for a strong serials collection in engineering, as well as extensive conference proceedings, especially in computer-related fields. Beyond that, selectors supporting doctoral research need to pay special attention to miscellaneous formats. As simple as the serial/monograph ratio is to implement, collections must go beyond books and journals and incorporate government documents, web resources, standards, datasets, and other gray literature. A research-level collection relies on diversity of format as well as diversity of idea and, if statistics is any indication, even a low serial/monograph ratio can still allow for extensive book and journal use. More broadly, the significant variation between the internal and external results indicates a need for librarians to study local research habits as well as external ones. At Mason, internal and external results can be averaged to generate tailor-made serial/monograph ratios. Other institutions, armed with their own internal data, could follow suit.

Age of Cited Resources
The age dispersion of STEM resources suggests that researchers—even in fast-paced, high-tech fields—are using resources from throughout the twentieth century. The oldest materials tend to be foundational works of historical importance to the field (such as a seminal patent or *On the Origin of Species*). In some cases, older materials lie outside the subject area and are used as sources of data or interdisciplinary support (for instance, a book of English and Scottish ballads used in a computer science data-mining project). These materials may be outside the collecting responsibility of the engineering librarian, but they are used by engineering researchers all the same.

FIGURE 2
Age of Resources Cited



*Internal (dissertation) citations.

Subfields that are more stable or well established (like statistics) or those with roots in historical data or preexisting fields (like operations research or civil engineering) show the most evenly dispersed age ranges. This trend is illustrated in figure 2: narrower data point groupings within civil engineering, operations research, and statistics reflect the researchers' preference for materials from a variety of publication dates, with less bias toward any particular timeframe. These subfields use current sources but have been developing long enough to have an older core of serials and monographs. (Correspondingly, many of these subfields show relatively high monograph usage.) While the internal results show a more even age dispersal overall (that is to say, even computer science shows a much tighter grouping than most of the external citation subfields), the same pattern is evident within the internal results, with SEOR and especially STAT more tightly clustered than other internal subfields.

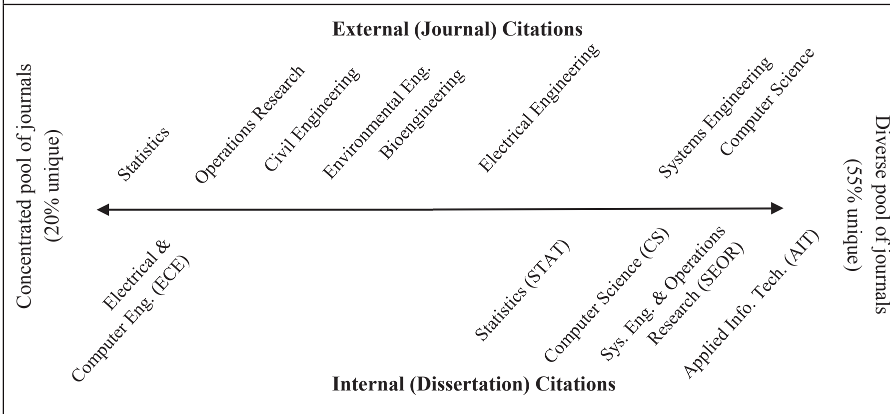
Computer-related fields (and bioengineering) show the opposite trend, skewing more decidedly toward recent publications. In figure 2, this is reflected in the more scattered grouping of data points; bioengineering, computer science, and electrical engineering are particularly broad, with a clear preference for materials from the last ten years. In bioengineering, the relative recentness of the subfield is particularly evident. Bioengineering researchers are not necessarily using the newest resources (less than 10 percent are from the last two years), but resources are heavily concentrated in the last 15–25 years. This assessment of bioengineering is supported by the low percentage of books cited, but it may change over time as the field develops further.

Compared to the external citations, dissertations rely on more recently published sources, especially in computer science, systems engineering, and operations research. Correspondingly, graduate and postgraduate students in these three subfields are more likely to be hindered by delayed access to materials. At the same time, scholars writing dissertations do still use older resources, covering the same broad range as the external citations. Again, this is reflective of a dissertation's exhaustiveness. In some cases, dissertations from a given subfield will even cite a higher number of older resources than their external counterparts. Both statistics and electrical engineering follow this pattern, reinforcing the idea that even in engineering we should not assume that all resources must be brand-new. For Mason—an institution that has expanded dramatically in the last 15 years—this means that retrospective collection development may be necessary to accumulate these older, still-cited materials. A concerted effort to fill in the last 26–50 years of engineering literature (for instance, by purchasing journal backfiles) could satisfy as much as 99 percent of the need.

Journal Distribution among Cited Resources

With regard to the number of unique journals cited in each subfield, two major groupings emerge from the results: subfields with very high numbers of unique journals and subfields with relatively few unique journals. Those subfields with larger numbers of unique journals include computer science and AIT—large, fast-growing fields with a

FIGURE 1
Diversity of Journals Cited, Based on the Percentage of Citations Indicating Unique Titles



plethora of niche interests. One only has to review the catalog of IEEE publications to understand why computer-related sciences might cite so many individual journals. Further, because AIT encompasses information technology across all disciplines (as do other areas of computer science), AIT research draws from many subjects. The other subfield with an especially large number of unique journals is systems engineering, which like AIT is interdisciplinary and thus warrants a large variety of resources. At the other end of the spectrum is statistics, a relatively focused field with a well-established core. The remaining subfields, such as civil engineering and environmental engineering, have broad reach but a definite core literature.

Journal distribution patterns among the internal citations vary slightly from the external patterns, but not beyond what might be expected owing to the unique habits and research interests of a local research community. ECE is one exception, showing a much smaller core of journals than in the external results, as well as broader overall usage of materials. Why electrical engineering varies so greatly from, say, computer science is unclear, but it is probably idiosyncratic to Mason. Statistics shows the inverse pattern, with internal citations from STAT revealing a much broader pool of journals than in the external group. This may reflect the need for dissertations to draw exhaustively and innovatively from a wide range of sources; it may also be the result of the study methodology. Nearly half of the STAT dissertations sampled were from the Computational Sciences and Informatics department; and, while they were classified with the subject heading "Statistics," they may nevertheless deviate from statistics research norms. Overall, these results suggest a need to be aware of local researchers' preferred core resources as well as the broader scholarly distribution. External-based title lists are insufficient, as are internal, usage-based lists on their own.

Additional collection development implications of the journal distributions are at least twofold. First, it appears that subfields with a broad spread of journals are fast-paced, diverse, and even interdisciplinary in nature. These areas should be monitored closely to see how patterns develop; more funding may be necessary to keep pace with publications, or it may be logical for several interdisciplinary subfields (such as computer science, AIT, and electrical engineering) to be funded cooperatively. Further, the distribution of journals highlights the need to look at more than the serial/monograph ratio when determining funding. Journals are purchased by the title, not the article, so the number of unique journals within a subject literature is just as important as the total number of citations. A 50:50 ratio means very different things when that "50" comprises 30—rather than, for instance, ten—unique journals. Many of these journals will be used lightly, and researchers can ILL them as necessary, but others may receive enough use to warrant subscriptions. Collection development will depend, as always, on a careful understanding of local need and attention to the broader standards.

Conclusion

This study set out with three major goals: First, it sought to ascertain whether internal and external citation analyses yield distinct results, even when addressing the same narrow STEM fields. In fact, the results did vary between internal and external sources; an internal analysis of statistics, for example, showed that 59 percent of resources used were journals, versus 75 percent came from external citations. More than 20 percent of internal systems engineering citations were for resources published within the last two years, versus just 5 percent among external results. In short, internal and external citation analyses tell very different stories. Any institution wanting a balanced collection—one that meets local need as well as external standards of excellence—cannot rely wholly on just one method of citation analysis.

The second purpose of this study was to satisfy the question of balance in citation analysis results by testing a method for combining external and internal citation data. This was also achieved. With results from both internal and external citation analysis, it was possible to determine averages for the appropriate age and format of a given subject collection, developing a more nuanced serial/monograph ratio than most other studies can provide. For example, rather than basing the Mason statistics collection on a serial/monograph ratio of 59:41 drawn wholly from internal data, or an external ratio of 75:25, librarians at Mason can aim for a ratio closer to 67:33—a custom blend of the external and internal results. Other libraries could easily average the 75:25 ratio with their own internal findings.

The third and most straightforward goal of this study was to provide usable information on the citation habits of STEM researchers. As predicted, variations do exist among subfields of a given discipline, and researchers in different settings do use materials distinctly. Journals remain a top format, but not without competition from conference proceedings and books, particularly in the computer sciences. Other subfields, like bioengineering, display stronger preference for journals. While engineering is generally thought of as a cutting-edge field, citations in all subfields include resources from the entire twentieth century. Computer science and electrical engineering cite the newest resources, while statistics and civil engineering cite more historical materials. Across all subjects, most citations refer to materials from the past 50 years. With regard to the diversity of materials cited, fields with the largest number of unique journals include computer science and systems engineering; statistics cites the smallest core of journal titles. Based on these variations, it becomes clear that there is no such thing as a one-size-fits-all citation analysis—not even within a given discipline. Based on these detailed findings, librarians at Mason can begin to sketch a profile of each subfield of engineering, using both Mason-specific and more universal data.

Ultimately, citation analysis—whether internal or external—yields useful and abundant data. That said, the method is not without its pitfalls and limitations, most of which have been explored in previous studies. Obvious limitations include the potential for bias in selecting sources; the amount of labor involved in collecting and annotating citations; and the challenge of classifying citations, particularly government documents and electronic resources. Citation analysis also has more complex problems: for instance, it only credits sources that are cited, rather than all the resources consulted during the research process.[30] It also overlooks other uses of the literature beyond publication, such as instruction or basic subject orientation.[31] Finally, citation analysis is based on the assumption that researchers are expert searchers, scouring the entire universe of literature exhaustively. In truth, this is rarely the case, and citation analysis results will inevitably exclude a host of resources that were simply overlooked. Thus, citation analysis can only be part of the collection development equation, bolstered by data from other sources.

Nevertheless, in spite of its shortcomings, citation analysis (and the serial/monograph ratio it yields) is a useful component of collection development, assessment, and budgeting. Citation analysis can help confirm or dispel assumptions about the research habits of scholars across the curriculum; it can also show differences between the internal and external research populations, provide a basis for collection priorities regarding age and format, and underscore the particular needs of doctoral-level researchers. In the end, many of the method's weaknesses stem from its labor-intensiveness and the fragmented nature of citation analysis efforts, which can be mitigated.

Moving forward, librarians (particularly those who enjoy membership in consortia) should implement the approach tested here, collaborating on external citation studies using standardized methodologies and a range of source materials. These studies should cover subfields at a specific and granular level without straying too far into local idiosyncrasy. Over time, the combined efforts of many institutions could create a finely tuned master list of globally derived serial/monograph ratios—a pool of external data gathered without undue burden on any one institution. At the same time, institutions should move forward with internal citation analysis to determine the realities of local need. Even brief citation analyses (1,000 citations or fewer) might shed light on unique habits of the institutional population and help shape collection development decisions for the better. Wherever possible, they should include faculty publications, rather than relying exclusively on student work. Then, armed with both internal and external data, librarians can develop their own appropriate serial/monograph ratios, collection ages, and core title lists. One size does not fit all; but, with a combination of collaborative external studies and small-scale local analysis, libraries can compile appropriately nuanced "characteristics of the literature" data without having to do everything themselves.

Notes

1. Robin B. Devin and Martha Kellogg, "The Serial/Monograph Ratio in Research Libraries: Budgeting in Light of Citation Studies," *College and Research Libraries* 51, no. 1 (Jan. 1, 1990): 46–54.

2. Paul L.K. Gross and E.M. Gross, "College Libraries and Chemical Education," *Science* 66, no. 1713 (Oct. 1927): 385–89.

3. Devin and Kellogg, "The Serial/Monograph Ratio in Research Libraries," 46–54.

4. Pali U. Kuruppu and Debra C. Moore, "Information Use by PhD Students in Agriculture and Biology: A Dissertation Citation Analysis," *portal: Libraries and the Academy* 8, no. 4 (Oct. 2008): 387–405.

5. Frances A. Delwiche and Ellen F. Hall, "Mapping the Literature of Athletic Training," *Journal of the Medical Library Association* 95, no. 2 (Apr. 2007): 195–201, doi:10.3163/1536-5050.95.2.195; Vincent Larivière, Éric Archambault, and Yves Gingras, "The Place of Serials in Referencing Practices: Comparing Natural Sciences and Engineering with Social Sciences and Humanities," *Journal of the American Society for Information Science and Technology* 57, no. 8 (June 2006): 997–1004, doi:10.1002/asi.20349.

6. Daniel P. Coffey, "A Discipline's Composition: A Citation Analysis of Composition Studies," *Journal of Academic Librarianship* 32, no. 2 (Mar. 2006): 155–65, doi:10.1016/j.acalib.2005.12.001; Jennifer Wolfe Thompson, "The Death of the Scholarly Monograph in the Humanities? Citation Patterns in Literary Scholarship," *Libri* 52, no. 3 (2002): 121–36.

7. Claudia Lascar and Loren D. Mendelsohn, "An Analysis of Journal Use by Structural Biologists with Applications for Journal Collection Development Decisions," *College and Research Libraries* 62, no. 5 (Sept. 2001): 422–33.

8. Clyde W. Holsapple et al., "A Citation Analysis of Business Computing Research Journals," *Information and Management* 25, no. 5 (Nov. 1993): 231–44, doi:10.1016/0378-7206(93)90072-2; Coffey, "A Discipline's Composition," 155–65; Helen Georgas and John Cullars, "A Citation Study of the Characteristics of the Linguistics Literature," *College and Research Libraries* 66, no. 6 (Nov. 2005): 496–515; Juris Dilevko and Keren Dali, "Improving Collection Development and Reference Services for Interdisciplinary Fields through Analysis of Citation Patterns: An Example Using Tourism Studies," *College and Research Libraries* 65, no. 3 (May 1, 2004): 216–41; Scott A. Curtis, "Informing Collection Development through Citation Examination of the Civil Engineering Research Literature" (paper presented, *ASEE Annual Conference and Exposition*, Vancouver, British Columbia, June 26–29, 2011); Delwiche and Hall, "Mapping the Literature of Athletic Training," 195; Li Zhang, "Citation Analysis for Collection Development: A Study of International Relations Journal Literature," *Library Collections, Acquisitions, and Technical Services* 31, no. 3/4 (2007): 195–207, doi:10.1016/j.lcats.2007.11.001.

9.Jeffrey H. Waller, "Evaluating Scholarly Communication at the Subdisciplinary Level," *Collection Management* 30, no. 2 (2006): 45–57, doi:10.1300/J105v30n02_04; Margaret Allen, Susan Kaplan Jacobs, and June R. Levy, "Mapping the Literature of Nursing: 1996–2000," *Journal of the Medical Library Association* 94, no. 2 (Apr. 2006): 206–20; Melissa L. Rethlefsen and Lisa C. Wallis, "Public Health Citation Patterns: An Analysis of the American Journal of Public Health, 2003–2005," *Journal of the Medical Library Association* 95, no. 4 (Oct. 2007): 408–15, doi:10.3163/1536-5050.95.4.408; Susan Kaplan Jacobs, "Mapping the Literature of Maternal-child/gynecologic Nursing," *Journal of the Medical Library Association: JMLA* 94, no. 2 (Apr. 2006, suppl.): E56–E64; Mary K. Taylor,

"Mapping the Literature of Medical-surgical Nursing," *Journal of the Medical Library Association: JMLA* 94, no. 2 (Apr. 2006, suppl.): E65–E73.

10. Thompson, "The Death of the Scholarly Monograph in the Humanities?" 121–36; Larivière, Archambault, and Gingras, "The Place of Serials in Referencing Practices," 997–1004; Jennifer E. Knievel and Charlene Kellsey, "Citation Analysis for Collection Development: A Comparative Study of Eight Humanities Fields," *Library Quarterly* 75, no. 2 (Apr. 1, 2005): 142–68, doi:10.1086/431331.

11. Erin T. Smith, "Assessing Collection Usefulness: An Investigation of Library Ownership of the Resources Graduate Students Use," *College and Research Libraries* 64, no. 5 (Sept. 2003): 344–55; Pierre Feyereisen and Anne Spoiden, "Can Local Citation Analysis of Master's and Doctoral Theses Help Decision-Making about the Management of the Collection of Periodicals? A Case Study in Psychology and Education Sciences," *Journal of Academic Librarianship* 35, no. 6 (Nov. 2009): 514–22, doi:10.1016/j.acalib.2009.08.018.

12. Kuruppu and Moore, "Information Use by PhD Students in Agriculture and Biology," 387–405; Laura Newton Miller, "Local Citation Analysis of Graduate Biology Theses: Collection Development Implications," *Issues in Science and Technology Librarianship* 64 (2011): 3, doi:10.5062/F4RF5RZZ; Graham Sherriff, "Information Use in History Research: A Citation Analysis of Master's Level Theses," *portal: Libraries and the Academy* 10, no. 2 (Apr. 2010): 165–83.

13. Melissa L. Rethlefsen, "Citation Analysis of Minnesota Department of Health Official Publications and Journal Articles: A Needs Assessment for the RN Barr Library," *Journal of the Medical Library Association* 95, no. 3 (July 2007): 260–66, doi:10.3163/1536-5050.95.3.260; Thomas Schaffer, "Psychology Citations Revisited: Behavioral Research in the Age of Electronic Resources," *Journal of Academic Librarianship* 30, no. 5 (Sept. 2004): 354–60, doi:10.1016/j.acalib.2004.06.009; Gloriana St. Clair and Rose Mary Magrill, "Undergraduate Use of Four Library Collections: Format and Age of Materials," *Collection Building* 11, no. 4 (Dec. 31, 1992): 2–15, doi:10.1108/eb023316.

14. St. Clair and Magrill, "Undergraduate Use of Four Library Collection," 2–15.

15. Schaffer, "Psychology Citations Revisited," 354–60; Miller, "Local Citation Analysis of Graduate Biology Theses," 3; Sherriff, "Information Use in History Research," 165–83; Laurel A. Haycock, "Citation Analysis of Education Dissertations for Collection Development," *Library Resources and Technical Services* 48, no. 2 (Apr. 2004): 102–06.

16. Feyereisen and Spoiden, "Can Local Citation Analysis of Master's and Doctoral Theses Help Decision-Making?" 514–22; Kuruppu and Moore, "Information Use by PhD Students in Agriculture and Biology," 387–405.

17. Smith, "Assessing Collection Usefulness," 344–55; St. Clair and Magrill, "Undergraduate Use of Four Library Collections," 2–15.

18. Devin and Kellogg, "The Serial/Monograph Ratio in Research Libraries," 46–54; Larivière, Archambault, and Gingras, "The Place of Serials in Referencing Practices," 997–1004; Curtis, "Informing Collection Development."

19. Holsapple et al., "A Citation Analysis of Business Computing Research Journals," 231–44.

20. St. Clair and Magrill, "Undergraduate Use of Four Library Collections," 2–15; Virginia Kay Williams and Christine Lea Fletcher, "Materials Used by Master's Students in Engineering and Implications for Collection Development: A Citation Analysis," *Issues in Science and Technology Librarianship* 45 (2006), doi:10.5062/F4PG1PPR; Patricia Kirkwood, "Using Engineering Theses and Dissertations to Inform Collection Development Decisions Especially in Civil Engineering" (paper presented, *ASEE Annual Conference and Exposition*, Austin, TX, June 14–17, 2009); Robin N. Sinn, "A Local Citation Analysis of Mathematical and Statistical Dissertations," *Science and Technology Libraries* 25, no. 4 (2005): 25–37, doi:10.1300/J122v25n04_03.

21. Georgas and Cullars, "A Citation Study of the Characteristics of the Linguistics Literature," 496–516; Zhang, "Citation Analysis for Collection Development," 195–207; Dilevko and Dali, "Improving Collection Development," 216–41.

22. Larivière, Archambault, and Gingras, "The Place of Serials in Referencing Practices," 997–1004.

23. Thompson, "The Death of the Scholarly Monograph in the Humanities?" 121–36; Smith, "Assessing Collection Usefulness," 344–55.

24. St. Clair and Magrill, "Undergraduate Use of Four Library Collections," 2–15.

25. Holsapple et al., "A Citation Analysis of Business Computing Research Journals," 231–44; Coffey, "A Discipline's Composition," 155–65; Lascar and Mendelsohn, "An Analysis of Journal Use by Structural Biologists," 422–33; Taylor, "Mapping the Literature of Medical-surgical Nursing," E65–E73.

26. Kirkwood, "Using Engineering Theses and Dissertations to Inform Collection Development Decisions"; Curtis, "Informing Collection Development."

27. "Facts & Figures: 2011–2012," George Mason University Libraries, available online at library.gmu.edu/about/ui/facts.php [accessed 20 November 2013].

28. Smith, "Assessing Collection Usefulness," 344–55; Miller, "Local Citation Analysis of Graduate Biology Theses," 3; Larivière, Archambault, and Gingras, "The Place of Serials in Referencing Practices," 997–1004; Curtis, "Informing Collection Development"; Kirkwood, "Using Engineering Theses and Dissertations to Inform Collection Development Decisions."

29. Sinn, "A Local Citation Analysis of Mathematical and Statistical Dissertations," 25–37.

30. Kuruppu and Moore, "Information Use by PhD Students in Agriculture and Biology," 387–405.

31. Eugene Wiemers Jr. et al., "Collection Evaluation: A Practical Guide to the Literature," *Library Acquisitions: Practice and Theory* 8, no. 1 (1984): 65–76, doi:10.1016/0364-6408(84)90055-3.