7-2-2008

Citation Analysis of Upper-Level Student Writing at Trinity University: Final Report

Benjamin R. Harris
Trinity University, bharris@trinity.edu

Follow this and additional works at: http://digitalcommons.trinity.edu/lib_faculty

Part of the Library and Information Science Commons

Repository Citation
Harris, Benjamin R., "Citation Analysis of Upper-Level Student Writing at Trinity University: Final Report" (2008). Library Faculty Research, 9.
http://digitalcommons.trinity.edu/lib_faculty/9

This Working Paper is brought to you for free and open access by the Coates Library at Digital Commons @ Trinity. It has been accepted for inclusion in Library Faculty Research by an authorized administrator of Digital Commons @ Trinity. For more information, please contact jcostanz@trinity.edu.
Introduction

During the 2006-07 academic terms, a citation analysis of third- and fourth-year student writing was conducted at Trinity University’s Elizabeth Huth Coates Library. This study was designed to achieve the following objectives:

- To improve the content focus and quality of library instruction at the Coates Library.
- To allow Trinity librarians to better support the goals and objectives of the university’s honor code.
- To offer useful information about student research and writing to interested Trinity faculty members.
- To offer suggestions for teaching research and writing strategies to upper level students.

During the course of the research, additional implications of value to librarians, teaching faculty members, and students were discovered. The following final report for this project offers a review of the results of this study, implications of the study for readers in the Trinity community, as well as implementation data for other information professionals.

Citation Analyses: Purpose and Process

Examples of citation analyses of essays written by students are somewhat common in the library literature. Consistently, the purpose of these studies is to evaluate a library’s collection policies. Every year, librarians at university’s across the country will spend millions of dollars on items they assume to be useful to the local academic community (Magrill and St. Clair 27). These decisions are often based on circulation data, trends in research topics, student and faculty requests, interlibrary loan data, and comparisons with other institutions. Data related to the actual use of sources by local community members is one strategy for evaluating collection methods that extends beyond a dependence on expert judgment (Wallace and Van Fleet 215). Conducted across libraries and institutions, these studies have also been used to develop “core lists” of journals in specific subject areas (Burright, Hahn, and Antonisse 200).
As with other evaluation methods, citation analysis is not an exact science in terms of collection evaluation. Students may use interlibrary loan in the course of their research, an activity that may be more common in one academic subject area as opposed to another (Haycock 103). Students may also cite items only to meet the criteria of an assignment, regardless of the value of this use (Waugh and Ruppel 277). Ultimately, without continued research to determine “why” a source was used, the researcher must be satisfied that a citation analysis can only record the number of times an item was cited. The value and the “why” of the cited material is more challenging to diagnose.

In the late 1990s, librarians and information professionals began to conduct citation analyses to gauge the impact of the World Wide Web on student research methods. While such studies have been able to chart changes in the use of sources based on method of delivery, the results are not surprising. The use of resources available on the web has led to a decrease in the use of scholarly materials in traditional formats (Davis and Cohen 309). However, many of these studies take the short-sighted view that web resources are inherently “unscholarly” and qualitative evaluations of these results rarely consider that electronic sources might be appropriate if not preferred depending on the assignment or subject matter (Kushkowski 262; Young and Ackerson 89).

A number of recent studies have explored citation analysis as a means to evaluate student learning and library instruction. While there are other methods for evaluating student learning upon receiving library instruction, the review of essays and bibliographies allows librarians the rare opportunity to quantify actual student behavior (Hovde 4). This type of review also allows the librarian to make connections between objectives developed for instruction sessions and the achievement of these objectives in student work. Ultimately, research designed for this purpose has been rare, possibly due to claims that intervening factors may create interference between the point of library instruction and the point of student activity (Young and Ackerson 174). In addition, comparative research on student learning would require a control group to determine that students would not be able to achieve the same results without library instruction, and it is rare for library instructors to refrain from teaching for the sole purpose of research (Mohler 58).

Outside of the library literature, citation analyses have been utilized for a number of other reasons. Some disciplines use this type of research to assist in making claims about the value of published material (its “citedness” or “citeability”), information that is often used in promotion and tenure decisions. In the linguistics literature specifically, claims about the traditional and changing “nature” of the academy as a discourse community are made via analysis into the ways sources are used or introduced in scholarly writing.

A recent trend in citation research has been to focus on the accuracy of citations and the credibility of citation structure. Due to concerns related to increases in plagiarism as well as intellectual property issues and copyright law, increased attention has been given to the accuracy of student citation practices. In addition, the transient character of web resources and difficulties involved in citing these sources has become the focus of some citation analysts since these problems can create barriers to access (Casserly and Bird
Still, inaccurate attribution is nothing new. Writers have created faulty bibliographies and failed to give proper attribution since the development of more complicated formats for citation in the mid-19th century (Sweetland 293-4). However, the province of training and assistance in citation method has shifted over time, from all disciplines, to writing courses, and now to libraries, giving these concerns relevance for librarians and other information professionals (Harris 4).

While the aims of citation analysts may differ, strategies used to code and record written work tends to be consistent. Most often, bibliographies are dissected to determine the form of source being used (book, journal, electronic resource, thesis/dissertation, other); at times, the specific titles of works are recorded. Publication dates and the scope of resources have also been analyzed. The primary difference between studies is the selection of material used by the researcher. In most cases, theses and dissertations or published works are selected for citation analyses research since they are readily available and permissions are rarely required for the study of these documents. In fewer cases, essays written by students during the course of a specific class are chosen. While this strategy requires greater effort on the part of the researcher and the consent of teachers and students, the results are likely to be more indicative of the kinds of writing conducted by students consistently during their matriculation. Theses and dissertations, while useful articles of study, are often produced only once and exhibit specific conventions of that writing form. The study of theses and dissertations may not prove useful to determine how students use sources in a day-to-day, semester-by-semester manner.

The methodology for this project was developed after a review of the professional and scholarly literature across several academic subjects (librarianship, English/Composition, linguistics). See the bibliography for a complete listing of works cited and consulted during the course of this study.

**Project Methodology**

**Recruitment**

At the beginning of the fall 2006 academic semester, faculty members across disciplines received an email from Coates Library director Diane J. Graves requesting volunteers to allow their classes to be included in this study. Those interested were asked to contact the project director, Benjamin R. Harris, Assistant Professor and Reference/ Instruction Librarian. If a department was not represented among those volunteering, the project director and other librarians sent requests to specific departments and faculty members.

After soliciting the campus for volunteers, faculty members in six academic departments allowed their students to participate in the project.
Procedure of Subject Participation

For classes that require submission of assignment in print form, faculty members were instructed to ask students to submit a second copy of their written work. The second copy was to be submitted in tandem with the version that would be graded by the professor. The second copy did not include identifying information such as name, student ID number, etc. Instead, to determine that all of the papers submitted were written by upper level (third- and fourth-year) students, participants were instructed to include their academic year on the cover sheet or first page of the document.

For classes requiring electronic submission of resources, electronic copies of essays excluding name identification were submitted to the library’s senior secretary via email. Documents were printed and routed to the project director. Since Microsoft Word files (the most popular form of word processing software for Trinity students) often records the names of individuals who receive a license for the software, this method of delivery was designed to maintain anonymity. In addition, email information was intercepted by a second party prior to the project director’s receipt of the documents.

The initial participation goal for this project was 150 submissions from across disciplines written during the fall 2006 semester. As of January 2007, 80 submissions were received. Two papers from a humanities discipline were rejected because the writer was a 2nd year student. Two papers, one from a humanities discipline and one from business, were excluded because the writers failed to include bibliographies with the submissions. One essay from the social sciences was excluded because the paper itself was incomplete. While the bibliography was intact, data related to the use of sources could not be collected. Another essay from the social sciences was excluded because the citation style was not recognizable or was not consistent enough to suggest that a specific style was in use. In total, six papers were rejected. 74 samples of assigned essays written by students were included in the study.

Data Recording

After receiving submissions from faculty members, papers were labeled to indicate academic department, starting with A, and continuing with AA after Z. This label was then used to organize papers into three categories of results: papers from humanities disciplines, papers from social sciences disciplines, and papers from the sciences.

After receiving all submissions, a fourth category was created for papers submitted from business courses. These papers were produced as part of a group project; these groups were comprised of target students as well as students in their first and second years at the university. This data was recorded and is included separately from the social sciences numbers. Due to the specific character of these texts, statistics in the business category should not be used in a comparative analysis.
Each student paper was then identified by number and year of submission. For example, the first paper for the first class submitted by a faculty member was coded 1A-2006. If necessary, this code could be used to designate information collected from a specific paper.

Citations on each writer’s bibliography were numbered, coded for type of material (Journal, Book, Electronic, Theses/Dissertation, Other), and then reviewed to determine the accuracy of the citation format. The essay was then read and each parenthetical citation (or footnote including attributive material) was numbered and coded based on the method of source integration (direct quotation, paraphrase, reference). Direct quotations were labeled as such regardless of whether the use was integral or non-integral to the sentence structure. The paraphrase category refers to both paraphrased statements and summary statements. Statistics for references include two methods of source use: (1) the writer refers to the source without including information taken from the source or (2) the writer includes multiple parenthetical citations for a single paraphrased statement.

Identification of Subjects in Stored Materials and Informed Consent of Participants

While student writers were made aware of their participation in the study, students could not be identified based on their submission to the project director. Therefore, informed consent was not necessary. Further, faculty members participating in the study could not be identified in the recorded data. Based upon exception #3 in Trinity University’s Institutional Review Board Procedures description, this project did not require formal provisions for informed consent. This was confirmed by the chair of Trinity’s IRB after a review of the research protocol for this study.

Quantitative Statistics

Data collected over the course of this research project have been organized into two sections. The first tables provide statistical averages for all of the data recorded during the coding process for each individual example of student writing. The second set of results is organized based on specific questions that were defined during the pre-research process.

It should be noted that among the papers studied, there were a number of writers who produced perfect bibliographies with no errors in the integration of their sources. In others, students were unable to create a single correct citation. These averages should be viewed as a general “snapshot” of current practice by student writers.
According to their essays, students in the humanities are twice as likely to choose books over other types of sources. In terms of citation quality, they are the least likely to offer an accurate or correct bibliography in comparison to other disciplines.

Students in the humanities are more likely to paraphrase as opposed to quoting sources. Rarely do these writers include citations only as references to work by other writers, and instead, use quotations or paraphrased passages taken directly from a selected text. The most common use of a research source is to explain or note an idea in support of the author’s thesis.

One sample in the 74 selected for this study failed to offer attribution for the use of sources. This sample from the Humanities category of essays failed to include citation information for the use of three direct quotations.

Journal articles are the most common type of source cited by students writing in the social sciences. Like writers in the humanities, they are less likely to use sources beyond the book and journal format.

Consistently, student writers in social sciences classes offer accurate bibliographic information at the end of their essays. The use of sources in the text of essays by these students was most commonly integrated in the form of paraphrased attributions to other literature. In particular, writers tended to cite specific numbers and facts in the course of their writing.

One social sciences sample included parenthetical citations for works that were not included on the bibliography at the end of the essay.
As one would expect due to the timely nature of scholarly literature in the sciences, writers in these courses depend heavily on journals. Books are rarely included in their bibliographies based on the findings of this study.

Considering the number of citations and the frequent use of outside sources in the science literature, bibliographies and in-text parenthetical citation of sources show a high level of accuracy.

In comparing the textual use of research, writers in the sciences include more citations than their counterparts in the other disciplines. While most of these appear in the form of paraphrased statements, students in science courses are the most likely to refer to sources that they do not cite directly. In general, the tendency to refer to sources with which the reader may be familiar, or to depend on the reader’s ability to seek out additional examples of research, is a characteristic specific to scholarly literature in the sciences.

The averages shown for students writing in business classes require one caveat. Papers included as a part of this citation analysis project focused on a current issue or topic. For this reason, a greater dependence can be seen on electronic sources, most commonly in the form of online news resources. In many cases, the same information could have been located via library resources or a scholarly database. Students appeared to prefer to cite the World Wide Web publications of news releases, as opposed to using the same sources as archived in the library’s holdings.

In general, bibliographies for these samples were the least consistent. Students used a number of styles; however a fair number of samples used different citation models for the same type of material. Two books might be cited in two completely different ways. It is
possible that since these were produced by groups of students, several different people
were involved in constructing the citations on these bibliographies. It was apparent that
they were not reviewed for consistency upon submission.

The lack of either publication date of material or access date for online resources was the
most common error in bibliographic citations. In the coding of the examples from
students writing in a business class, web links were analyzed for a publication date. Only
those that failed to include a publication date where a publication date was clearly
presented are counted under the column of “Incorrect” citations. Citations for electronic
resources that failed to include an access date were considered incorrect without further
review.

Due to the varied length of papers in this sample, only the first five pages were analyzed
for their use of in-text references to outside literature. With this in mind, the frequency of
elements of cited material in business would likely compare to the use of sources in the
sciences.

Qualitative Results

The research protocol for this citation analysis specified four questions to be answered
over the course of the study. These questions are listed below in their original form,
followed by answers and additional commentary.

1. Of the sources included in bibliographies for essays written by third and fourth year
Trinity students during the 2006-2007 academic year, how many are (1) library print
resources, (2) library electronic resources, (3) Internet resources, and (4) other resources.

   By a fair margin, journals see greater use in student work across disciplines when
   compared to research materials available in other formats. While the statistical averages
   above do not distinguish between articles accessed electronically and print journal usage,
   electronic access is dominant (at an estimated rate of 20 to 1).

   Books are rarely cited and those that are included are often included in the assigned
   reading for the class. This can be deduced when a number of papers from a class use an
   identical print source.

   It may surprise some readers to learn that Internet resources tend to be used only when
   the assignment itself calls for information that students expect to find on the World Wide
   Web only (current news).

   Few “other” resources were included in bibliographies. This category includes items in a
   format that would not be considered a book, journal, or web resource (such as a pamphlet
   or lecture notes).

   During the data gathering stage, researchers were prepared to find theses or dissertations
   among the items cited. Similar studies showed at least a limited number of citations for
   these materials. None of the samples in this study included citations for theses and
dissertations.
2. Of the citations included in bibliographies for essays written by third and fourth year Trinity students during the 2006-2007 academic year, how many of the citations are constructed correctly according to an established (assigned) style manual?

Save for a number of exceptions, the bibliographies analyzed during this study suggest that Trinity students are competent in terms of citation construction and in-text attribution. The highest rate of error occurs for World Wide Web resources. However, these errors may have less to do with a student’s ability to create citations. Due to the lack of a standard for the presentation of information about titles, authors, publication dates, etc. on internet sites, the challenges involved in locating the information required to create a citation will be likely to have an impact on the accuracy of the citation.

In addition, style manuals offering citation formats for online resources are often complex and confusing. Writers may not understand what information should be included or how best to organize this information based on the samples provided in a particular style manual. Writing handbooks are even more problematic (due to the brief treatment given to the specificity of citation styles for varying formats) and students using these resources are further disadvantaged.

3. Of the in-text references to citations included in student essays written by third and fourth year Trinity students during the 2006-2007 academic year, how many of the in-text references are constructed correctly according to an established (assigned) style manual?

Overwhelmingly, in-text references created by students were accurate and consistent. One of the reasons for this may be that in-text references are consistent across different resource formats. The in-text format for a book is the same for an electronic source, and so forth.

4. How is information taken from cited sources used in students’ writing? (Primarily, is the use of cited sources lateral or hierarchical in purpose? Lateral citation includes cited fact or opinion used to support the claims of the student writer; he or she agrees with the source. Hierarchical citation includes cited fact or opinion that the student writer seeks to discredit or expand beyond; he or she does not agree with the source or feels the source is deficient in some manner.)

Based on the results of this study, Trinity students are most likely to paraphrase an original source when integrating the material in their writing. Paraphrasing suggests a personal synthesis as well as an integration into the prose of an essay, both of which are positive attributes in student writing.

Almost exclusively, information from sources came in the form of facts. Statistics and research results were the most common types of information cited in this study’s sample. Students are most likely to adopt a lateral use of the source, where the facts and information located are used to support the ideas of the author. They are rarely used in a hierarchical manner, where the writer questions the source or the information provided.
Implications of Results

The following observations based on the research process and the results of this study are organized into three groups. This should not suggest that the implications of results for one category will not be of interest to readers in other categories.

For Students

Often, student writers are asked to join in the discourse of a particular community or to mimic the communication styles of various groups. Strategies and modes of communication are not always explicit, and students will often require guidance and training to write as community members. (See David Bartholomae’s “Inventing the University” for an extended explication of these activities.)

Models of writing by scholars and students are often used to guide writers as they attempt to gauge the requirements of communicating within specific communities. While students were not the intended audience of this report, a comparative review of results from this study might allow them to develop a sense of how students research and write for particular disciplines.

For Teaching Faculty

Students are using journal resources at a greater rate when compared to the use of other types of sources. In many ways, this is a shift from the time when access and use of material published in books was the easiest or most efficient method for students gathering information. Faculty members may need to encourage the use of books when they believe these sources will be necessary or more useful in relation to a specific assignment or topic. In addition, the common critique that students are using more electronic resources rather than traditional journal articles or books cannot be supported by the results of this study.

Studies in linguistics and academic discourse, as well as research in library studies, show that there are very different perceptions and expectations of scholarly writing in different communities. This is a timely reminder that big general claims about “scholarly writing” should be carefully scrutinized. Writing in one discipline varies with the writing in other academic areas, and this is also true of expectations for student writing. It may be a surprise for someone teaching in one discipline to see student work from different classes or disciplines, and the results of studies such as this offer access to this information.

In terms of fostering good citation practice and the ethical use of sources, the researcher’s reading and analysis suggest that instructors should assign a citation style. While encouraging students to choose their own citation style seems productive, papers from courses without a consistent citation style were more problematic than those with a selected style. While some will argue that students may never cite another source upon graduating from the university, they will continue to be asked to follow directions…to fill
in the blanks. It is not without purpose to suggest that some rigor be involved in the
citation of sources.

For Librarians

Librarians designing learning tools and teaching information literacy sessions may
benefit most from research into the products of student work. Most often, librarians must
make assumptions or hypotheses concerning the results of instruction and similar
activities, as they are seen in the product of student work.

Research studies such as this allow liaison librarians for specific academic subjects to
develop better focus for instruction sessions. For example, liaison librarians in the
sciences may learn that their students will need expertise in searching the journal
literature, since the written discourse of that discipline relies heavily on the use of
research using journal resources. A librarian in the humanities may notice that students
need greater assistance when citing sources, possibly due to the preferred citation style of
that course/discipline. A sound pedagogical motivation in planning instruction for
students is to determine “where students are,” and while this can rarely be determined
within the scope of a single 50-minute instruction session, a sense of “where students
were” is a better source of information than none.

Conclusion

Two factors, one practical concern and one issue in the methodology, challenge the
usefulness of these results. First, the limited number of volunteers for the classes makes it
difficult to draw comparisons between academic areas. In relation to the goal of
maintaining anonymity for participants, small sample sizes limit the collection of
statistics related to specific disciplines and make it easier to track or trace the course. In
order to fashion a more robust and malleable set of results, greater participation in the
form of faculty and student volunteers would be necessary.

In addition, the desire to determine quality of “use” related to resources—a peripheral
goal of the study—is challenging in light of specific integration and research usage
practices among disciplines. Ultimately, the results of this type of qualitative analysis
could not suffer a comparison, since the scholarly literature on linguistic practices
between academic discourse communities suggests that such claims would be futile (if
not impossible to construct).

In future, if librarians or others wish to analyze primary written documents by university
students, they should seek to secure an ample set of volunteers in order to make “across
the university” results possible. Otherwise, researchers may wish to limit their analysis
to core classes within specific disciplines to provide a more focused and/or detailed
snapshot of writing by undergraduates. This would also provide an appropriate situation
for qualitative research into the use of sources, with respect to the traditions and
aesthetics of that particular discourse community.
Works Cited and Consulted


Casserly, Mary F., and James E. Bird. “Web Citation Availability: Analysis and Implications for Scholarship.” College and Research Libraries 64.4 (July 2003) 300-16.


Dunkin, Paul S. “Happiness is a Long Footnote.” Library Resources and Technical Services 7.4 (Fall 1963): 403-5.
Harris, Benjamin R. “Credit Where Credit is Due: Considering Ethics, Ethos, and Process in Library Instruction on Attribution.” *Education Libraries* 28.1 (Summer 2006): 4-11.


Magrill, Rose Mary, and Gloriana St. Clair. “Graduate Term Paper Citation Patterns By Disciplines and Level of Course.” *Collection Management* 12.3-4 (1990): 25-56.

Mohler, Beth A. “Citation Analysis as an Assessment Tool.” *Science and Technology Libraries* 25.4 (2005): 57-64.


Richmond, Phyllis A. “Misery is a Short Footnote.” *Library Resources and Technical Services* 9.2 (Spring 1965): 221-4.


