Commercial Database Design vs.
Library Terminology Comprehension:
Why Do Students Print Abstracts
Instead of Full-Text Articles?

Bonnie Imler and Michelle Eichelberger

When asked to print the full text of an article, many undergraduate college students print the abstract instead of the full text. This study seeks to determine the underlying cause(s) of this confusion. In this quantitative study, participants (n=40) performed five usability tasks to assess ease of use and usefulness of five commercial library databases and were surveyed on their understanding of library terminology. The study revealed that more than half of the students correctly defined the term “Abstract” and over 75 percent understood “full text.” However, only 25 percent of the students were able to successfully complete all five database tasks.

Over the past 10 to 15 years, while working at the reference desk, the authors have come across hundreds of printouts of journal article abstracts lying unclaimed next to the library’s printers. These orphan abstracts beg the question, “Do students know the difference between an abstract and the full text of an article before they hit the print button?” For all of the abstracts that were left unclaimed, the authors knew that many more were being given to professors as examples of journal articles and being used as sources for research papers. In a previous research study, the authors found that, of 39 students who had been exposed to the concepts of “abstract” and “full text” through library instruction, only 62 percent were able to find and print the full text of five articles related to their research assignment. The remaining 38 percent of students printed at least one abstract in place of the full text.1

The authors were interested in studying the underlying cause or causes of this disconnect. Did our students not know the difference between an abstract and the full text of an article, even when they had received instruction on these concepts? Or were the database results pages designed so that the full text of the article was too difficult to find? Or was there a combination of factors leading to the students’ obvious confusion?

The authors selected Proquest as the test database for their previous study because it was the most popular database at Penn State University, and it was the only aggregate database listed on the “Try These

---

Bonnie Imler is Library Director at Penn State Altoona; e-mail: bbi1@psu.edu. Michelle Eichelberger is Systems and Electronic Services Librarian at Genesee Community College; e-mail: maeichelberger@genesee.edu. © 2014 Bonnie Imler and Michelle Eichelberger, Attribution-NonCommercial (http://creativecommons.org/licenses/by-nc/3.0/) CC BY-NC
First” research help page. While Proquest includes a high percentage of full-text articles, the full text of some articles can only be retrieved by using SFX citation linking software to connect to more specialized databases, which are less familiar to undergraduates. Any student who uses SFX to retrieve the full text of a Proquest citation will run into databases whose format and design are very different from that of Proquest. The authors watched as students struggled to find the full text of the articles in these databases repeatedly during the first study and noted that participants had an especially difficult time finding the full text of articles in the following five databases: JAMA, Springerlink, Oxford Journals, Cambridge, and Pediatrics (Official Journal of the American Academy of Pediatrics).2

The authors prepared a follow-up study to determine if the failure to find the full-text articles on these citation/abstract pages was due to the student failure to comprehend the difference between “abstract” and “full text,” or if it was due to the design of the databases’ web pages. This follow-up study is relevant because, even though most Penn State students do not start their research process with these specialized databases, they are likely to encounter these and other more obscure resources when they use the Penn State citation linking service to retrieve the full text of a citation-only article from Proquest or other general aggregate database.

The authors designed a test scenario to determine if database design was an influence on the full-text discovery success rate. The authors followed the online testing activity with a survey on the terms “abstract,” “full text,” and “pdf,” to discover how well the participating students understood the database terminology. Analysis of the terminology survey and the screen capture video from student research sessions showed that, in the case of the five databases selected, database design was more of a deterrent to task completion than student misunderstanding of library terminology.

**Literature Review**

Library jargon has been a problem since long before the development of online resources. According to Vaughn and Callicott, “Library terminology has consistently been a “sticky wicket” for librarians and library users…”3 Naismith and Stein’s article about library jargon from 1989 reported that patrons misunderstood library terminology approximately half of the time.4 John Kupersmith keeps a running list of terms reported by libraries as being misunderstood on his website.5 Spivey’s study from the year 2000 showed that potentially unclear library terminology appeared on all 60 of the college and university library home pages included in the study.6 A review of the literature shows that many of the problems that occur during library website usability testing have more to do with patron failure to comprehend library terms such as “database, periodical, or catalog” than with poor web design.7 Kruegar et al. points out that “The majority of studies agree that users are frustrated by confusing library terminology and an overwhelming amount of information.”8 Understanding that library jargon could affect student ability to find the full text of articles from the five selected databases, the authors designed this study to include a survey to test student comprehension of the three most vital words needed for article retrieval: “abstract,” “full text,” and “pdf.”

The usability study is a common and well-respected tool used for the assessment of academic library web pages, and the literature contains dozens of articles on the practices and issues of usability testing, most of them “how-to” studies and descriptions. Usability studies may be implemented in response to results from library evaluation tools like LibQUAL, or in preparation for a major site upgrade.9 They are also used as ongoing assessment tools for libraries.10 Usability study methodology can include reviewing website and database usage logs, focus groups, direct observation, card sort protocol, and think-aloud protocol.11 Most academic library usability
studies encompass the entire library website, covering usage of the library catalog, its informational pages, and information related to periodicals and databases. Many of these studies included one or two tasks designed to encourage participants to use the library’s databases, and several studies focused specifically on the library page portion of the article retrieval process, but no studies have focused on the usefulness of web design of the databases themselves.

A note should be made here about the difference between “ease of use” and “usefulness” in terms of usability studies. Vaughn and Callicott point out the risks involved in usability studies in their article “Broccoli Librarianship and Google-Bred Patrons.” Library users may be more attracted to web pages that are as easy to use as a Google search box, but focusing on the ease of use of a library website may cause the researcher to lose sight of the “usefulness” or the added value of library material over a simple Google search. Paraphrasing researcher Stanley Dicks, Vaughn and Callicott note that “usefulness refers to the overall usefulness of the product. Does it do what it is supposed to do? Is it usable at all? Does it work? [as opposed to simply being easy to use].” Tsakonas and Papatheodorou had a slightly different perspective on the “ease of use” vs. “usefulness” debate, and considered ease of use to be a crucial part of usability, which they defined as focusing on “the effective, efficient and satisfactory task accomplishment and aims to support a normal and uninterrupted interaction between the user and the system.” The authors of this study were interested in both ease of use and usefulness, but they designed the study to focus primarily on whether or not the five selected databases “worked,” meaning that they were useful in leading the student researcher to the full text of an article.

Methodology
Participants
After receiving approval from the Penn State University Institutional Review Board, the authors began recruiting students for the study. In the previous study on usage of the SFX citation linking software, the authors recruited students who had already received library instruction, with the hope that most of them would have heard of the SFX service during their instruction class and that they would be able to complete the tasks assigned to them by using SFX. For the follow-up study, the authors were curious to see if native understanding of the terms “abstract” and “full text” made any difference in the ability of students to successfully find the full text of an article in the five selected databases. To test if the five selected databases were usable without any special training, the researchers recruited first-year students who had never received library instruction during their university experience. Jannik and Whang et al. have noted that library instruction can affect usability testing results, and the authors wanted to remove that variable from the study.

After contacting faculty members who do not regularly request library instruction, the lead author was invited into eleven 100-level courses to present her research topic and to recruit participants. The author stressed that participation was voluntary and in no way influenced students’ course grades. While planning the study, the authors knew that, according to Jakob Nielson, only five students were needed for a qualitative usability study, but they were interested in completing a quantitative study to generalize broader user behavior, which, according to Nielsen, required twenty students. Hoping to gather even more quantitative data, the authors capped the student participation number at forty.

Forty undergraduate first-year students over the age of seventeen were self-selected for the study. Each participant was given an appointment card marked with a time and date on which to arrive at the lead author’s office. No additional demographic information was captured for the study. Students were given an implied
informed consent form that described the study and any risks involved. A statement on the form indicated that completion of the computer-based research session implied that the student consented to take part in the study.

**Materials**
A basic Dell computer loaded with the screen capture software TechSmith Morae was used to collect and analyze the data in this study. This computer was similar to the desktop models available to students in the library, and the default Internet browser Internet Explorer was also familiar to the participants from their use of campus computers. The authors chose to use Morae as their screen capture tool because it gave them the ability to set up discrete tasks/assignments for the participants, it allowed the participants to leave the canned assignment to perform live interaction with preloaded websites, and because it could also be set up to capture survey data. Like other screen capture software such as SnapzPro and Camtasia, Morae can be set up to record audio, video, on-screen activity, and keyboard/mouse strokes during a defined period of time.

The study included an electronic survey, activated by the Morae software after each participant completed the five research tasks. The survey asked the participants to define, in their own words, the terms “abstract,” “full text,” and “pdf.” The authors chose to administer the survey after the completion of their research tasks so that the students would not be alerted to the importance of those terms in their research tasks, which focused on finding the location of the full text of an article on a database citation page.

**Research Design and Procedure**
Batteson et al. note that “in formal usability testing, users are observed using a site, or prototype, to perform given tasks or achieve a set of defined goals.”

For this usability study, the authors set up one task to be completed in five different commercial databases. Study participants were asked to find and print the full text of five preselected articles (one per database) from the citation/abstract page in the following five databases subscribed to by Penn State University: JAMA, SpringerLink, Oxford Journals, Cambridge, and Pediatrics (Official Journal of the American Academy of Pediatrics). These five databases had proved difficult for students to navigate in the authors’ previous study on SFX usage in Proquest, and the authors were curious to see if student participants in the follow-up study would face similar usability challenges. While more obscure than Proquest or other popular aggregate databases, these five databases are commonly accessed via the aggregate databases by students who use the SFX citation linking software.

While SFX enables students to link to the full text of articles in other database platforms, it does not always open the full text of the article for the student researcher. Instead, the link resolver often leads the researcher to the citation/abstract page rather than to the full text of the article. It was observed in the previous study that students who successfully followed the SFX links from Proquest to the citation/abstract page of a different database were then unsuccessful in identifying the location of the full text of the article, often marked by the words “Full Text” or a PDF symbol. The authors designed this study to test student ability to locate the full text of the article from the citation/abstract page (that is to say, the SFX final landing page) in the five databases listed above.

The authors designed a welcome screen in Morae that provided the participants with instructions for navigating through a series of five tasks. This welcome screen appeared as a small gray box centered at the top of the computer screen. Instructions for each subsequent task were presented in a similar gray box. The task box contained a clickable URL designed to take the participant directly to the citation page of each database. (See figure 1.) The tasks were set up to be
self-initiated and were not time-limited. When the participant started a task, the gray box would shrink, so as not to be a distraction. A small “show instructions” button in the shrunken gray box allowed the participants to review the instructions if necessary. An “Exit Session” button gave the participant the opportunity to exit the study at any time.

Because the focus of the study was participant interaction with database search results pages, and not their database search strategies, the authors decided to create direct links to each database citation/abstract page. This saved the participants’ time and made the data analysis more efficient because the authors did not have to ignore or delete the extra screen capture time that it would have taken the participants to search for an article in each database. It also made the data more consistent because it negated the chance of participant typos and poor navigation, and it made it easy for every participant to interact with the same screens in the same databases. The authors chose the citation/abstract database page as their link destination because it is the page most often linked to when using citation linking software to find the full text of an article.

Participants were asked to print the articles so that the researchers could identify the moment in time when the students thought that they had completed the task and found the full text of the article. Without forcing the student to print, the researchers would have had to guess as to when the student thought that he or she had completed the task and found the full text. Using a “faux printing” process developed in a previous study24 to avoid potential printer technology problems and to save paper, the computer was set to send the articles to a nonexistent printer. On the welcome screen, participants were given the following instruction: “The articles you print will be sent to a printer outside of the room. After selecting print, you may move on to the next task.”

All forty of the participants used the same computer for this study. To avoid having previous research session activities affect the following sessions, the authors set the “visited link” color in the Internet browser to be the same color as the “nonvisited link” color so that students would not be alerted to links that had been tried by other participants. The authors also deleted the print cache after each session so that the next participant could not compare their printing choices with those of the previous participant.

**Data Collection**

On his or her scheduled research day, the study participant met with the lead author in the library, where he or she was taken to a quiet study room with a single personal computer on a desk. The participant was given a copy of the implied informed consent form to read and was encouraged to keep it in the event that he or she needed to make contact with the investigators in the future. The research consent form informed the student that screen capture software was in use and that it would be recording movements on the screen.

![FIGURE 1 Task Box](image)
The student was seated at the computer, where the lead author activated the welcome screen for the study. The lead author then left the room, and the participant began the five assigned tasks. Screen capture began when the participant clicked the Start button on the welcome screen. Upon completion of the five tasks, the library terminology survey appeared on the computer screen. After the survey questions were completed, a thank-you message appeared, followed by a button that prompted the participant to open the study room door to seek the researcher. The researcher then assigned a number to the file and saved it to the hard drive. When the student completed the session, he or she was given a $5.00 gift certificate for a local favorite convenience store located near the campus.

Data Analysis
After all of the student research sessions were completed, the authors used Morae to auto-compile task completion times and responses to the vocabulary definition section. The lead author then examined each participant recording and coded it to note when the participant completed each task, and whether he or she printed the abstract of the article or the full-text article (html or pdf format). As well as indicating if a particular action had occurred, each marker also noted the moment in time when that action happened so that the researchers could look at how slowly or how quickly each participant completed each task and at which point during the task the action occurred. As the lead author reviewed the files, she created ad hoc markers to note unexpected actions taken by the participants, such as printing articles outside the scope of the task. The authors used Morae to compile the marker data and exported it into Excel spreadsheets for review.

Findings
Database Tasks
Out of the forty students, only ten (25%) successfully printed the full text of all five articles from the five different databases. Forty percent of the students printed all abstracts, and 35 percent printed 1–4 abstracts. Of the sixteen students who printed abstracts only, six of them clicked on print immediately after the URL loaded for all five articles and did not spend any time scrolling or looking for a full-text indicator. The average time to complete all five tasks and the survey question was 7.63 minutes. The shortest amount of time spent was 4.27 minutes and the longest was 12.95 minutes. (See figure 2.)

![Student Completion of Research Tasks](image-url)

**FIGURE 2**
Student Completion of Research Tasks
Survey Results
Twenty-one, or slightly more than half of the forty participants, were able to successfully define the term “Abstract” as related to library research. Correct student definitions included “Preview of text or synopsis,” “a summary of a research article,” and “a cut-down version of the full document.” An even higher percentage of the participants (31 of 40) understood the concept of full text, defining it as “to view the entire text,” and “the whole article not just the abstract and conclusion,” and “the entire reading.” Pdf was more difficult for the students to define in their own words. Many students guessed at defining the acronym (“published document file?” “preferred document format,” “print document file”), while others admitted that they did not know what the acronym stood for, but they understood that it was an electronic copy of the original print article. A generous interpretation of the responses, including vague responses such as “the file that the article is saved to,” resulted in twenty-two correct answers from the forty participants. Accurate responses included “the original article,” “a type of file that represents the article exactly how it looks,” and “it is opened with Adobe and contains the full article.” (See figure 3.)

Discussion
Database Tasks
Like researchers Whang and Ring, the authors focused on task completion “as the primary evaluation method for measuring success.” According to this measure, only 25 percent of participants in this study were successful. Sixty percent of the students found the full text of at least one of the five articles, but that is a much lower success rate than the authors would hope for when Penn State Altoona students use the college’s databases to do research. Six of sixteen participants who did not find any full-text articles appeared to rush through the assignments and did not spend any time looking for a link to the full text. It is possible that they knew the difference between the full text and the abstract of each article, but they were more interested in receiving their $5.00 gift card than in spending time to thoroughly complete the assignments.

Comparison of the citation/abstract pages of the five different databases revealed the following web design flaws and inconsistencies that could have contributed to student failure to retrieve the full-text articles:

Inconsistent Terminology
The five databases used various means to identify the link to the full text of the article, often using several different
methods on the same page, including the words “full text,” “html,” and “pdf.” JAMA, Oxford, & Pediatrics had very similar interfaces, but used slightly different terminology for their links. JAMA used the terms “Full Text” and “Full Text (PDF),” while Oxford used “Full Text (HTML)” and “Full Text (PDF),” and Pediatrics used “Full Text Free” and “Full Text (PDF) Free.” Springerlink employed “Download PDF,” while Cambridge used the Adobe PDF icon next to the words “View PDF” and a globe icon next to “View HTML.”

From analysis of the mouse movements of the study participants, none of these terms appeared to trigger immediate recognition. Out of the eight students who completed the tasks correctly, only one went directly to the full-text link purposefully without scrolling the entire length of the web page to locate the link. The JAMA database did offer one additional point of access to the full text that the other articles did not. The text of the abstract of the article was followed by a plain text link reading “Full text of this article.” (See figure 4.) Participants seemed to respond better to this wording and location on the page, as eight students found and used this link.

**Location, Location, Location**

JAMA, Oxford, and Pediatrics all located their full-text links to the right of the article abstract. Springerlink and Cambridge placed their links above the abstract. All five of these databases positioned the links so that they’d be visible to a pc/mac.

---

**FIGURE 4**

JAMA Screen

![Image of JAMA Screen](image-url)
user in the initial screen load, regardless of the window size. However, 88 percent of the study participants did not see these links at first glance and scrolled down the page for at least one of the five tasks. Fifteen percent of the participants did not see or identify any of the full-text links at first glance, scrolling down the page in all five databases. It is interesting that the results of this study run contrary to those of Cockrell and Jayne, who noted that the students in their study were loath to scroll down the page to look for additional information.26 The researchers noted that, after scrolling down the pages, several participants hovered the mouse over the social bookmarking links, as if they were looking for the full-text link from within the Facebook, Twitter, and Del.icio.us icons. It is possible that the participants were drawn to the colors on the social media icons, which stood out in comparison to the more neutral color scheme of the rest of the citation/abstract result pages. (See figure 5.)
Font Size and Link Placement
When designing content for web pages, common practice is to put the most important information in the largest font, the next most important content in slightly smaller font, and so on. The authors found that, in all five databases, the largest font on the page was used for the title of the journal article, which makes sense because that is the best identifier for the article. However, since the main goal of the database vendor should be to provide access to the database content, the authors believe that the link to the full text of the journal article should have equal importance to the title, or only slightly less importance, and therefore should be created and displayed in a large font. In all five databases, the font of the links to the full text of the articles were smaller than that of the article title and were closer in size to that of the abstract text. The other problem with the full-text links on all of the database results pages was that the links were intermingled with other links and text and did not stand out clearly on the page, as seen in this screenshot from the Oxford Journals database. (See figure 6.)

The Tease
The SpringerLink database offered an additional feature the other databases did not, and this “feature” led to confusion and task failure for seven of the forty participants, or 17.5 percent of the students. An image of the first page of the article appeared directly under the abstract with the words Fulltext Preview above it. (See figure 7.) The image was formatted as a link, which seemed unnecessary, as the whole page was already visible on the citation/abstract page. Eight of the research participants clicked on the image, probably hoping to find the full text of the article, which was a logical assumption given that the word “fulltext” appeared in the image title. Clicking on the link caused a new window to pop up, but it did not provide the user with any additional content beyond the image that the user had already seen in the full-text preview window. Seven of the study participants were misled by this window and printed it out, thinking that it was the full-text view of the article. This database feature could be improved by redirecting the link to go to the full text of the article rather than to the first page of the article.

Survey Discussion
The authors were pleased with the results of the survey and were impressed with the native terminology understanding of first-year undergraduate students who had not received library instruction. More than half of the students were able to define all three research terms, and more than 75 percent understood the meaning of “full text.” This study’s 52.5 percent rate of comprehension for the term “abstract” was significantly higher than the percentage reported by Hutcherson in his 2004 study of library jargon, which was 36.20 percent of 297 first- and second-year undergraduates at California State University.27

A few of the answers given by the participants were more polished than the researchers would have suspected, leading to the suspicion that several participants may have used the pc’s Internet access to help them define the terms. However, even if these participants did use the Internet to help them define the terms, they chose the definitions that were most apt for the context, and the authors were satisfied that the students who listed the correct answers did have a good understanding of the concepts as related to library research.

After review of the results for the survey question about the definition of “pdf,” the authors realized that they should have designed the survey to ask the participants to define the word in relation to library research. Several of the respondents simply tried to clarify the acronym. About half of the students seemed to have a general idea of what a pdf would give them in terms of research, but the authors may have gotten a higher
percentage of correct answers if the question had been reworded and expanded.

**Correlation between Database Tasks and Survey**

The survey showed that more than half of the participants understood all three of the terms that they encountered in their preceding series of research tasks. More than 75 percent of the students understood what full text meant, which implies to the authors that if they had been able to find the link to the full text on each database citation/abstract page, they would have had a much higher success rate for the series of tasks than 25 percent. This is a significant finding, and it speaks to the inadequacy of the web page design of the five databases that were reviewed in this study.

**Limitations of Study and Suggestions for Future Research**

The literature reveals few examinations of databases in terms of research usability and ease of use, though there have been several studies of database usability in terms of adaptive software and their compatibility with screen-reading programs. Librarians and library staff have very little control over the appearance and ease of use of commercial databases.
and have focused their studies on web pages and websites where they can easily make improvements. Another reason to avoid focusing on commercial database usability is that as soon as a researcher identifies a problem in a particular database, it may be fixed in the next database update. Database vendors do perform their own usability studies, and some of them even employ UX (User Experience) Managers, but their studies may focus on their primary audience, often career professionals, and not the needs of undergraduate students. While there are limits and complications inherent in testing commercial databases for usability, the authors decided that the study results could be generalized to point out trends in commercial database development.

Allowing students to self-select for this study introduces the possibility of self-selection bias in the study results. Students who agreed to participate could have been more interested in, or familiar with, library procedures than the other students who did not choose to participate, which would skew the data toward a higher level of library terminology understanding than that of the general population. However, the high level of library term comprehension did not help these students succeed at finding the full text of articles in the five databases, so a case can still be made that both library terminology incomprehension and poor web page design can lead to article retrieval failure. On the other hand, students who self-selected only to receive the survey incentive, and who did not take their tasks seriously, could have skewed the results in a negative direction. However, the results still show a higher than suspected comprehension of the library terms “abstract,” “full text,” and “pdf,” and the six students who seemed to rush through the database tasks would only have made a 15 percent difference in the database task results if they had taken more time and been more successful, shifting the results from 25 percent success rate to a 40 percent success rate, which is still failing.

Screen capture video analysis is not an exact science. The authors have made some assumptions about why study participants may have been confused by the less familiar, more obscure database citation/abstract screens. A follow-up study could include focus groups where students would be asked to provide web page design feedback, such as where they thought that they would find the full text of the article on the database results list, or what they would change about the pages to make them more user-friendly. It is possible that this research could be approved and underwritten by the database vendors in a collaborative effort that could provide valuable data for their design teams.

Conclusions
The authors were pleasantly surprised to see that more than half of the study participants were able to define all three terms listed in the post-task survey, even without having attended a library instruction class at Penn State Altoona. Fully three quarters (75%) of the participants were able to correctly define the term “full text,” which does not correspond with the fact that only 25 percent of the participants were able to find the full text of an article in all five databases. The authors’ conclusion is that the participants understood what they were looking for but couldn’t find it on the databases’ citation/abstract screens, or were unable to correctly identify the pages that they found as full text or not full text.

What can be done about this? Librarians can spend more time teaching college students how to find the full text of articles in the library’s databases and focus on giving students the tools to identify the parts of a full-text article in any database that they might encounter during their research. Students may require more extensive instruction on how to identify an abstract and how it differs from the full text of an article. They may understand the theory of how these two views of an article are different, but they
may have trouble putting that theory into practice. As we move farther and farther from the world of print periodicals, today’s students have little sense of what an article should look like, and thus they may not be equipped to suspect that the article’s abstract is not the whole article in its full-text version.

More important, librarians need to do a better job communicating with database vendors, working with them to provide feedback to make the database web pages more user-friendly for students at all levels of college, even for beginning freshmen. According to usability expert Steven Krug, “The first immutable law of usability is, ‘Don’t make me think!’” Translated to the library environment, this means that librarians should not have to teach users to navigate their sites per se; rather, sites should be intuitive, and the location of needed resources should be easily identified by end users. Database vendors need to do a better job of making their websites intuitive, with standardized full-text article terminology, clear and obvious links to the full text of the articles, and site designs that are clean and accessible for all user levels, from first-year freshmen to graduate students. Librarians need to improve their communication with vendors to help them achieve the goal of making their sites more user-friendly for all academic researchers.

Notes
2. Ibid.
12. Whang and Ring, “A Student-Focused Usability Study”; Elizabeth Stephan, Daisy T.


23. Ibid.

24. Ibid.

25. Whang and Ring, “A Student-Focused Usability Study,” 82.

26. Cockrell and Jayne, “How Do I Find an Article?”


