Routledge
Taylor & Francis Group

Predatory Journals in Library Databases: How Much Should We Worry?

NERISSA NELSON AND JENNIFER HUFFMAN
University Library, University of Wisconsin–Stevens Point, Stevens Point, Wisconsin, USA

Much of the recent library literature related to scholarly communication and predatory publishers has focused on faculty concerns regarding publishing in questionable journals for tenure or promotion purposes. However, little attention has been paid to predatory publishers in the context of student research and library instruction. The presence of predatory journals in library databases may put students at risk of including questionable content in their academic output. While the results of this study reveal that the number of predatory publishers and their associated journals are fairly small in the three article database packages and one directory that were examined, predatory journal content was more prevalent in one particular resource and in certain subject areas.

KEYWORDS   Beall's List, journal databases, library instruction, Open Access publishing, peer review, predatory journals, predatory publishers

BACKGROUND AND INTRODUCTION

In April 2014, Jeffrey Beall, a librarian at the University of Colorado–Denver who writes extensively about the topic of predatory publishers, was invited to speak at the University of Wisconsin–Stevens Point's University Library as part of the Library's *Faculty Scholarship and Creative Works* series. His lecture, *Writers Beware! Predatory Publishers: A Serials Crisis for the 21st*

Century, prompted some questions from the authors of this article, specifically with regard to instruction and predatory journals.

Librarians and course instructors steer students toward high quality information sources to use in research papers and assignments. For articles, the most common strategy is to require students to use peer-reviewed journals. Most library databases and discovery tools (e.g., the Ex Libris product Primo) provide "peer reviewed" filters, where students can limit their search results to scholarly articles. During instruction sessions, and depending on the course and assignment, most librarians do not dedicate too much time—in what is typically a "one-shot" session—to explaining variations in quality among peer-reviewed journals. Although course instructors and librarians realize that the stringency by which the peer reviewed process is applied varies somewhat among journals, there is some degree of acceptance that even "lower tiered" academic journals available in the "peer-reviewed" section of databases and discovery tools provide a level of sufficient quality content for undergraduate research.

In 2009, Jeffrey Beall began creating a list of what he called "predatory publishers" after he discovered, through spam e-mails, a number of Open Access publishers who were corrupting or exploiting what is known as the "author pays" model, where the author pays a fee to the journal or publisher. Based on his discovery, he began a blog, "Scholarly Open Access: A Critical Analysis of Scholarly Open-Access Publishing," which tracks predatory publishers in an ongoing list. This list has grown exponentially since its inception and is often consulted by librarians and researchers.

This newer category of what has become known in the literature as "predatory publishers" and "predatory journals" claim to be both scholarly and peer reviewed. "Predatory" is a term that has become more common in the literature as a result of Beall's work. These types of publishers and journals have been identified by Beall to be of low-quality, driven by making a profit, and do not practice a real peer review process. Beall created criteria to identify these types of publishers and journals, which is described more thoroughly on his blog. For the purpose of this study, the authors have chosen to use the phrases "predatory publishers" and "predatory journals" throughout the article since our data collection and analysis is based on Beall's list of predatory publishers and journals. His list is currently the only comprehensive compilation of predatory publishers and journals that we are aware of and thus we chose to use Beall's list as a major tool for our research. Our study examines the extent to which these publishers and their journals, as identified by Beall, are present in select academic databases and in a directory commonly utilized in libraries.

With the emergence of predatory publishers and predatory journals, the assumption that the peer review claim provides a sufficient level of quality for undergraduate research should be revisited. The solicitation of qualified experts for the review, and the review itself, takes time—if it is done in a meaningful way. Quick turn-around between article submission and publication alone renders the peer-review claim of predatory journals questionable. In like manner it has

become evident that we need to examine the general assumption that the content providers libraries depend on have carefully examined and filtered content prior to its adoption in standard products offered to libraries. There has been a strong element of trust among librarians who purchase databases from vendors. Vendors often market their databases by highlighting the number of journals in particular disciplines, but they may not have a mechanism in place for weeding out the predatory publishers that have made their way—to date, modestly—into these databases alongside more trusted publishers.

For teaching librarians, the emergence of predatory journals raises a number of questions and issues. First, instructors need to know to what extent predatory journals are included in library databases. Is their presence negligible at this point in time, or has it become a serious concern? Do databases vary with regard to their inclusion of predatory journals? Is the presence of predatory journals particularly strong in certain subject areas? If the answers to the above questions give rise to serious concern, then a second cluster of issues has to be addressed: what are good strategies for addressing the distinction between truly peer-reviewed journals and predatory journals in library instruction? Do librarians need to alert course instructors about predatory publishers, and what would be effective ways of doing so? What can librarians, particularly those who oversee library database subscriptions, do to urge vendors to be alert to this problem and to carefully filter out suspicious content? Should librarians proactively manage predatory journal titles found within databases? These were the guiding questions for our research.

PREDATORY PUBLISHERS AND JOURNALS

The body of recent literature regarding predatory publishers and journals has focused mainly on the practices of these types of journals; how they are structured and managed through "editorial boards," and the questionable business models they use to make money.[1] Questions about current open access models, the peer-review process, journal prestige or impact factors, and the economics of information have been debated. However, little attention has been given to the extent to which content in predatory journals may be used by students in their own research papers and assignments. The following review of the literature provides background of what is known about predatory publishers, identifies gaps seen in the literature, and provides a framework for our study.

OPEN ACCESS MOVEMENT

The Open Access (OA) movement, as it emerged in the early 1990s, created conditions for predatory publishers to exist. The movement was seen as a

departure from the restrictive model of traditional subscription-based scholarly publishing to an environment where articles could be available more freely online with few or no restrictions, opening the world of research to a larger audience. Early groups and initiatives, such as the Budapest Open Access Initiative[2] and the Bethesda Statement on Open Access Publishing,[3] created policies and guidelines for how an author's work could be disseminated and distributed in this new environment. For example, the Bethesda's definition of OA requires two conditions for inclusion: (1) "The author(s) and copyright holder(s) grant(s) to all users a free, irrevocable, worldwide, perpetual right of access to, and a license to copy, use, distribute, transmit and display the work publicly and to make and distribute derivative works, in any digital medium for any responsible purpose, subject to proper attribution of authorship, as well as the right to make small numbers of printed copies for their personal use" and (2) "A complete version of the work and all supplemental materials, including a copy of the permission as stated above, in a suitable standard electronic format is deposited immediately upon initial publication in at least one online repository that is supported by an academic institution, scholarly society, government agency, or other well-established organization that seeks to enable open access, unrestricted distribution, interoperability, and long-term archiving (for the biomedical sciences, PubMed Central is such a repository)."[4]

There are different basic models within the OA movement, and variants within these models. The "Green OA" model is one in which authors self-archive a version of their work through an institutional repository (often after —or simultaneously with—publishing the work in a traditional journal), thus making the research accessible to the public and not "closed" behind traditionally restricted subscription access. Another model is called the "Gold OA" wherein authors publish their work in OA journals. Although the practice is most frequently associated with "Gold OA," under each of these models publishers may require authors to pay an article processing charge (APC). These fees charged back to the author go toward production/administrative costs.

The shift to OA is also taking hold as new policies are implemented. For example, The Higher Education Funding Council for England (HEFCE) has created a new policy effective April 2016 that states "certain research outputs should be made open-access to be eligible for submission to the next Research Excellence Framework (REF). This requirement will apply to journal articles and conference proceedings accepted for publication after 1 April 2016."[5] A similar OA policy will also take effect in the fall of 2015 as the U.S. Centers for Disease Control and Prevention and the Agency for Healthcare Research and Quality state that researchers will lose "…grant support from those sources if they don't make their findings freely available to the public."[6]

There are supporters and critics of the OA movement with gradations of opinion in between. Proponents laud OA as breaking down barriers to scholarly communication, and say that the academic publishing world should support OA to make research accessible to all, as a public service.[7] Critics of certain models of OA find that it has too great of a negative impact on scholarly communication where it leaves the door open to substandard journals with little or no quality control to take advantage of the eagerness of researchers to achieve publication.[8] Scholars, both new and established, have sometimes been vulnerable to predatory publishers' "bait the hook" scheme when seeking publishing venues or being asked to serve on editorial boards.[9] Some scholars have submitted legitimate research to these predatory journals while unaware of the true nature of the publishing operation due to the fact they appear to be authentic. Trying to retract those articles or having names removed from editorial boards has also proved to be a very difficult process.[10] These predatory operations tend to publish substandard or low-quality research through journal identities that mimic real and established journals by name and design, but are characterized by unrealistically fast turnaround times and inattention to quality and sometimes formatting issues. At the same time, it should be clearly stated that there are numerous legitimate and well-regarded OA academic journals whose aim is to remove access barriers to scholarly research. These journals do practice peer review and are highly regarded.

The Public Library of Science, more commonly known as PLOS, is an example of one of the first OA scientific publications that began in 2003 with *PLOS Biology*. After a letter from scientists to scientific and medical publishers fell on deaf ears at that time, the scientists behind PLOS established their own online journals, and the coalition became an OA publisher.[11] While a review process of submitted manuscripts exists in the PLOS environment, typically through an academic editor, and in some cases an external reviewer depending on the decision of the academic editor, acceptance rates of PLOS articles are significantly higher than those of traditional journals: 65–70% in comparison to 10–15%.[12]

Beyond the aim to open up research more freely and in a timely manner, OA proponents also intend to help curb increasing subscription costs that libraries can ill afford with typically stagnant budgets.[13] Totosy de Zepetnek and Jia suggest that "… high quality open-access journals will become established with prestige as universities realize the importance of publishing their own open-access journals in order to improve their own brand image and thus reduce the cost of the exorbitant subscription fees by large publishers because as subscriptions increase, libraries will continue to cut back. As a result of these various trends, open-access sources will garner wider circulation and become an important component in any academic's toolkit going forward."[14]

PEER REVIEW AND QUALITY CONTROL

The purpose of peer review in scholarly communication is to assess the quality and validity of research. This process ensures that research has undergone rigorous vetting wherein other experts in a given field evaluate presented research for accuracy, clarity, validity, methodology, and other criteria deemed important. The gold standard of peer review is the double-blind review, where neither authors nor reviewers have knowledge of each other. The merits of this process have been questioned for its supposed elitism, among other concerns.[15] Some suggest the process should not be left exclusively to formal reviewers or experts in the field, but should be expanded to the entire "readership community" to decide what is and is not quality research.[16]

For faculty, peer-review scholarship is one of the hallmarks of the academy. As a means of sharing knowledge with those in their discipline, it is also an important part of the tenure process. Dudley expresses concern that universities have created a scholarship problem and that they need to broaden their scope of research and consider alternatives beyond the traditional scholarly publishing model.[17]

Predatory journals and some of the articles published within them, may actually satisfy promotion requirements at some institutions and in some countries, such as Nigeria, where publishing research in an international journal is required.[18] Those who serve on tenure and promotion committees at universities in some instances may not be cognizant that it is a predatory journal title listed on a dossier. In the absence of some form of quality of the peer-review process, scholars who submit good research in a predatory journal may "lose face" if it is discovered that they did not recognize the journal to be suspect. Low quality work produced by scholars, or simply bad research that is accepted by predatory publishers also "loses scholarly relevance" and scholarship suffers in the process.[19]

SCHOLARLY PUBLISHING, LITERACY, AND LIBRARY INSTRUCTION

Teaching librarians know that most students and scholars, especially younger ones, prefer the online environment for browsing, searching, and retrieving information. They also know that students tend to favor Google, Google Scholar, and OA environments for their research. In addition to content available in library subscription databases and authenticated platforms, there is more research freely available today through OA platforms and repositories than has ever existed before.

Because of this change in scholarly communication, Warren and Duckett make a case that our traditional way of how librarians teach information literacy has "… rendered ineffective the instructional strategy of teaching students that 'library' equals 'good information' and 'free web' equals 'untrustworthy information.'"[20] They purport we need to teach students about academic publishing so they understand the complexities and economic factors of information, the "behind the scenes" of how technology creates access to free and fee-based information, and to go beyond the "mechanics and procedures" of retrieving information to how scholarly information is "created, vetted, and accessed" in today's environment.[21]

Warren and Duckett claim that scholarly communication awareness and information literacy should be integral for teaching librarians and will be one of the key factors in aiding in the growth of the OA movement in the future.[22] As McMillan and MacKenzie suggest, "shouldn't we, as instructional librarians, be concerned about students' abilities to use the information they have discovered?"[23]

With this in mind, Jeffrey Beall introduced "scholarly publishing literacy," to include skills such as understanding copyright and licensing agreements, bibliometrics, and indicators (a quantitative analysis of academic literature), how journals are ranked, publishing operations and mechanics, and other implications and issues of OA models.[24]

In her article "Riding the Wave of Open Access: Providing Library Research Support for Scholarly Publishing Literacy," Zhao discusses a situation wherein a librarian was helping a Ph.D. student with a manuscript he sought to publish. An e-mail the student had received had alerted him to a call for papers. Since the e-mail looked legitimate, provided an International Standard Serial Number (ISSN) for the journal, and mentioned where the journal was indexed, the student submitted his manuscript to the journal. After two weeks, his manuscript was accepted and he was notified that there would be an APC of $350. When the article came out a month later, the student discovered numerous errors and formatting issues. He contacted the publisher, but never received a response. The student then contacted a librarian at his campus for help, and the librarian found some startling problems with the publication. It was not indexed in the databases where it was supposed to be, nor was the journal found in the traditional periodical directories librarians use, such as *Uhlrich's Periodicals Directory* and the *Directory of Open Access Journals (DOAJ)*. The librarian also learned that the ISSN was fake, a common finding among predatory journals.[25] This example strongly illustrates a need to develop scholarly publishing literacy among both students and researchers.

For librarians, the issue of predatory journals presents a problem for those in charge of collections when trying to sift and winnow what is and is not valid through aggregators and vendors.[26] It is also problematic if predatory

journals are found and indexed in library databases because students may then use potentially bad research for their papers or assignments. The same holds true for researchers who may stumble on an article that looks valid to support their own work. One of the roles of librarians is to help patrons find the information they need or the best research that is available and to explain how to recognize and avoid low-quality journals, and at the same time make patrons aware of the misconceptions that are sometimes attached to legitimate OA journals.[27]

BEALL'S LIST

Jeffrey Beall was really the first to begin tracking predatory publishers back in 2009. His list of predatory publishers and standalone journal titles is updated and maintained on his blog, *Scholarly Open Access* (scholarlyoa.com), and has received more attention from librarians and scholars as more and more predatory publishers make their way into the OA world. The number of predatory publishers on Beall's list has increased from 242 in December 2012 to 772 as of April 2015.[28] His list has also garnered attention from academics who are concerned about their careers and the quality of publishing outlets where they submit their work.[29]

Many of the predatory publishers used the PLOS model as a framework, where hundreds of journals would be published under a single publisher name.[30] Also, journal titles with excessively broad topics, such as the "International Journal of Science and Advanced Technology" or "International Journal of Nuts and Related Sciences" exist under this foray and exploit the author-pays publishing model by making profits without necessarily delivering a quality product in return.[31]

The other problem with these low-quality journals is that they are often not preserved in the online environment or available in perpetuity, which would typically be an important factor for libraries as "keepers" of the record. Frequently, predatory journal content will disappear. Access to prior and cumulative research is important for the preservation or indexing aspect of content.[32]

It should be noted that Beall is not without critics. Crawford, for example, has found Beall to be Eurocentric by favoring Western publishers, and opines that he has a bias toward the major publishing house Elsevier. Other critics, such as Berger and Cirasella, who discuss the benefits and limitations of Beall's list of "potential, possible or probable" publishers, find that Beall often overlooks the fact that lower-tiered and even reputable academic journals in traditional publishing firms have on occasion published articles that are erroneous or questionable.[33]

While the composition of Beall's list may leave room for some criticism, his list, as mentioned earlier, is the most authoritative compilation of predatory publishers and journals available, so that we chose to utilize it in this study.

METHODOLOGY

Although the literature does point to a concern with the rise in predatory publishers and a lack of knowledge on the part of some scholars in recognizing them, no studies in the literature specifically examined how much content in predatory journals students may use in their own research papers and assignments as they retrieve articles from library databases.

The research questions for this study centers around this issue in hopes of better understanding how librarians should handle the problem of predatory content in the context of instruction, helping faculty with research, e-journal management, interlibrary loan, and vendor negotiations.

We examined three databases and one directory/index for predatory content. These databases—*ProQuest Central, EBSCO Academic Search Complete*, and *Gale Academic OneFile*—were chosen because they were represented as sources of high quality, scholarly full-text content available for multidisciplinary research in academic libraries. The index examined was the *Directory of Open Access Journals*, commonly known as DOAJ. DOAJ was chosen for comparison because it similarly advertised itself as an online directory that provided OA to quality multidisciplinary, peer-reviewed journal content. A brief description of the databases, *Beall's List*, and the index are provided below. Table 1 identifies quantitative data retrieved from all four resources.

Academic OneFile

Gale's website describes the *Academic OneFile* database as the "premier source for peer-reviewed, full-text articles for academic libraries" with comprehensive content from the "physical and social sciences, technology, medicine, engineering, the arts, technology, literature and many other subjects."[34] Based on the listing of publishers and titles obtained from the Gale website for this study, *Academic OneFile* contained content from 3,758 publishers with 16,555 titles covered by 753 identified subjects.

TABLE 1 Database/index publisher, title, and subject characteristics

Database or Index Name	Number of Publishers	Number of Titles	Number of Subjects
Academic OneFile	3,758	16,555	753
Academic Search Complete	3,801	13,787	1715
Beall's List	560	n/a	n/a
DOAJ	5,456	9,709	252
ProQuest Central	5,693	21,174	512

Academic Search Complete

According to EBSCO's description, the *Academic Search Complete* database is "the largest academic database available with the most full-text journals and broad depth of coverage" and it "provides complete coverage of multidisciplinary academic journals."[35] Based on the listing of journal publishers and titles obtained from the EBSCO website for this paper, *Academic Search Complete* provided access to content from 3,801 publishers representing 13,787 titles linked to 1,715 identified subjects.

Beall's List

Beall's List is described as a "list of questionable, scholarly open-access publishers."[36] *Beall's List* is part of Jeffrey's Beall's blog, *Scholarly Open Access*. At the time of retrieval of *Beall's List* for this research, it identified 560 publishers; however, the titles or subject areas associated with the list of publishers were not identified. The criteria used by Beall to determine if a publisher should be included on *Beall's List* is identified at https://scholarlyoa. files.wordpress.com/2015/01/criteria-2015.pdf.

DOAJ Index

DOAJ is the acronym for the Directory of Open Access Journals. DOAJ does not host any of the content that it lists; rather it is "an online directory that indexes and provides access to quality open access, peer-reviewed journals."[37] It functions as an index and gateway to full-text journal content that is available freely online. Based on the publisher and title list obtained for this research, the DOAJ indexed data represented 5,456 publishers with 9,709 titles covered by 252 identified subjects. For subject coverage, DOAJ states that publishers use the Library of Congress Classification outline.[38]

ProQuest Central

On their webpage, ProQuest described the *ProQuest Central* database as "the ultimate cross-disciplinary research tool" because it "brings together 30 of our most highly used databases to create the largest single academic research resource available today."[39] Based on the listing representing journal publisher and title data obtained from ProQuest, *ProQuest Central* listed content from 5,693 publishers with 21,174 titles represented by 512 identified subjects.

We set out to answer the following four research questions:

1. In each of the databases and the DOAJ Index, how many publishers matched with entries on *Beall's List*?
2. How many of the publishers on *Beall's List* showed up in more than one database or index?

3. In each of the databases and the index, how many journal titles were included from publishers identified on *Beall's List*?
4. In which subject areas were the predatory journals most prevalent?

DATA COLLECTION

The data collection portion of the research consisted of gathering information from various sources and creating lists of publisher names, journal titles, and other associated data. The sources of these data included Beall's predatory publisher list, vendor database publisher/title lists (EBSCO's *Academic Search Complete*, Gale's *Academic OneFile*, and ProQuest's *ProQuest Central*), and the DOAJ index publisher/title list.

The list considered to be made up of predatory open access publishers was obtained from Jeffrey Beall's blog *Scholarly Open Access*.[40] The list was accessed in May 2014 and transferred to a Microsoft Excel spreadsheet for data analysis. The list included publisher names and Uniform Resource Locators (URLs).

The lists for the EBSCO, Gale, and DOAJ databases were downloaded from their respective websites, while ProQuest's list was obtained from the vendor.[41] The lists for EBSCO, ProQuest, and DOAJ were downloaded in May 2014 while the list for Gale was obtained in September 2014.

For the vendor databases and DOAJ index, the original lists were acquired either as .csv or .xls files and managed for analysis using Microsoft's Excel software program. When available in the lists, the primary fields that were examined and used for comparison purposes included the following:

- Publisher
- Journal title
- Alternate title
- ISSN
- EISSN
- Subject
- URL

Since the publisher list available from Beall's blog only contained publisher names and URLs, other sources of information were needed to cross-reference and verify if a publisher or journal title was or was not the same as listed in the database and DOAJ index lists. Additional sources of information that were consulted included web pages of the journals as identified on *Beall's List*, the vendor or index lists, and *UlrichsWeb*, an international periodical directory.

DATA ANALYSIS—PUBLISHER NAMES

After all of the data was collected, a working list in the form of an Excel spreadsheet was created to manage publisher names. The purpose of the working list was to compare the publisher names from the predatory publisher list to the names used in the databases and DOAJ index. The working list included all of the publisher names from the predatory publisher list, the three databases, and the DOAJ index.

When working with the DOAJ list, it became clear that publisher names were not uniform and a publisher may be entered in the list several times using variant forms of the name. For example, the publisher *Academy & Industry Research Collaboration Center* was listed two times using the variant forms *Academy & Industry Research Collaboration Center (AIRCC)* and *AIRCC*. As another example, the publisher *Internet Scientific Publications, LLC* was listed three times using variant forms of the name including *Internet Scientific Publications LLC; Internet Scientific Publications, LLC*; and *Internet Scientific Publications, Texas*. Because of variant forms of one publisher name throughout the DOAJ list, attempts were made to deduplicate variant forms of the same publisher and clean up the data. However, with over 5,600 publisher entries, it is not certain that every variant form of a publisher name was removed.

Because of the lack of uniform spelling and naming convention of the publisher names and the inconsistent presence of ISSNs and EISSNs, Excel's "find" feature was used to manually search the working list for publisher name matches across all of the data. Using a keyword searching strategy, manual searching was performed across the publisher names used in *Beall's List*, the databases, and the DOAJ index. The keyword search strategy incorporated the publisher name used in *Beall's List* or other variants of the name. If a direct match was found between the predatory publisher name and the database or index publisher names, that name was flagged in the working list. If a reasonable match was made but was not exact, the other data attributes of a publisher (e.g. ISSN, URL, journal titles) were used to verify if the name used represented the same publisher. If the variantly named publisher was found to be the same one, that publisher name was also flagged. When the search for all publishers across all of the database and index fields was finished, the number of matches was summed and the discovered percentage of predatory publishers in each resource was computed.

DATA ANALYSIS—JOURNAL TITLES

The journal titles associated with the predatory publishers in the databases and the DOAJ index were identified from the original lists available from EBSCO, ProQuest, Gale, and DOAJ. To determine the subject areas of the journals offered by the predatory publishers in the databases and DOAJ index,

working lists were created to manage the publishers and journal titles using Microsoft's Excel software. A separate list was created and managed for each database and index. Each list was identified by the vendors' or index's name and included the publisher name and their corresponding journal titles.

After the journal titles were identified, and because each vendor had a different naming convention for the subject area of their content, the journal titles were categorized for this research using a simplified list of subject areas. Initially, the following 14 general subject areas were developed to classify the journal titles for each predatory publisher in a database or the DOAJ index:

- Arts or Entertainment
- Business
- Communication
- Education
- Government or Politics
- History
- Humanities or Social Science
- Language or Literature
- Medicine or Health
- Multidisciplinary
- Philosophy or Religion
- Science
- Sports or Recreation
- Technology

Several factors were considered when classifying the journal titles including the publisher's own classification, *UlrichsWeb*'s classification, and the researcher's interpretation based on a review of the journal title on the publisher's website. When a journal title's subject area was not identified or could not otherwise be determined, a subject area was not assigned. When several subject areas could be applied to the journal's content, the subject area was designated as Multidisciplinary.

Based on the generalized subject areas assigned to the journal titles associated with a predatory publisher's journal title in the databases or the DOAJ index, the subject areas were summed and a percentage of journal titles per subject area were calculated for each of the predatory publishers in the databases and DOAJ index.

RESULTS

This section presents the results of the comparisons between *Beall's List* of publishers and the publishers and journal titles in three academic databases and one online journal index.

TABLE 2 Percentage of predatory publishers by databases and index

Database or Index Name	Total Number of Predatory Publishers	Total Number of Publishers	Percent Predatory Publishers (%)
Academic OneFile	2	3,758	0.05
Academic Search Complete	6	3,801	0.16
DOAJ	123	5,456	2.25
ProQuest Central	41	5,693	0.72

Research Question 1: In each of the databases and the *DOAJ* Index, how many publishers were on *Beall's List*?

After comparison of the publisher names in each database with *Beall's List* relative to the total number of publishers in the index, the DOAJ contained the highest percentage of the products we examined, at 2.25%. In DOAJ, there were 123 publishers matching up with entries in Beall's list out of the 5,456 total publishers shown. The *ProQuest Central* database contained the second highest number of these publishers at 0.72% (41 out of 5,693), followed by *Academic Search Complete* at 0.16% (6 out of 3,801) and *Academic OneFile* at 0.05% (2 out of 3,758). Table 2 lists the quantitative results of the predatory publisher listings in each of the databases and the DOAJ index.

Research Question 2: How many of the publishers on *Beall's List* showed up in more than one database or index?

None of the predatory publishers with content in *Academic OneFile* were listed in any other database or index. In *Academic Search Complete*, four out of six (66.6%) of these publishers were listed in either the DOAJ index or *ProQuest Central*. In the DOAJ index, 28 out of 123 (22.8%) of these publishers were listed in either *Academic Search Complete* or *ProQuest Central*. And in *ProQuest Central*, 28 out of 41 (68.3%) of these publishers were also listed in either *Academic Search Complete* or the DOAJ index. Table 3 identifies the publishers on *Beall's List* that were listed in more than one database or index.

The content for two of the publishers on *Beall's List* (Advanced Research Journals and Southern Cross) was found in three of the databases and DOAJ index including *Academic Search Complete*, DOAJ, and *ProQuest Central*.

Research Question 3: In each of the databases and the index, how many journal titles were included from publishers identified on *Beall's List*?

TABLE 3 Predatory publisher listings in multiple databases/index

Academic Search Complete	DOAJ	ProQuest Central
Advanced Research Journals	Academic and Business Research Institute	Academic and Business Research Institute (AABRI)
MDPI Publishing	Academic Journals	Academic Journals
Natural Sciences Publishing Corporation	Academic Journals Inc.	Academic Journals Inc.
Southern Cross Journals	Academy & Industry Research Collaboration Center (AIRCC)	Academy & Industry Research Collaboration Center (AIRCC)
	Advanced Research Journals	Advanced Research Journals
	Asian Economic and Social Society	Asian Economic and Social Society
	AstonJournals	AstonJournals
	Bioinfo Publications	Bioinfo Publications
	Canadian Center of Science and Education	Canadian Center of Science and Education
	Canadian Research & Development Center of Sciences and Cultures	Canadian Research & Development Center of Sciences and Cultures
	Clute Institute	Center for Innovations in Business & Management Practice
	EconJournals	Clute Institute for Academic Research
	Human Resource Management Academic Research Society	EconJournals
	IACSIT Press	Human Resource Management Academic Research Society
	IBIMA Publishing	IACSIT Press
	Information Engineering Research Institute, USA	IBIMA Publishing LLC
	Institute of Electronic & Information Technology	Information Engineering Research Institute, USA
	International Foundation for Research and Development (IFRD)	Institute of Electronic & Information Technology
	Macrothink Institute	International Foundation for Research and Development
	MDPI	Macrothink Institute Inc.
	Mediterranean Center of Social and Educational Research (MCSER)	Mediterranean Center of Social & Educational Research
	Scholarlink Resource Centre	Natural Sciences Publishing Corp
	Sciedu Press	Scholarlink Resource Centre Ltd
	Science Publications	Sciedu Press
	SCIENCEDOMAIN International	Science Publications
	Scientific Research Publishing	SCIENCEDOMAIN International
	Southern Cross Publishing	Scientific Research Publishing
	The Center for Innovations in Business & Management Practice	Southern Cross Publisher

Note: None of the predatory publishers associated with *Academic OneFile* were identified in the other three information sources.

TABLE 4 Percentages of predatory journals in databases/index

Database or Index Name	Number of Journal Titles Associated with Predatory Publishers	Number of Journal Titles in Database or Index	Percent of Journal Titles Associated with Predatory Publishers (%)
Academic OneFile	6	16,555	0.04
Academic Search Complete	55	13,787	0.40
DOAJ	812	9,709	8.36
ProQuest Central	299	21,174	1.41

After the predatory publishers were identified in each of the databases and index, the journals titles associated with these publishers were identified. Of the four resources, the DOAJ index provided access to the most journal titles associated with these publishers. In DOAJ, a total of 812 journals titles, or 8.36% of its database listings, were identified as being available from the 123 of these publishers matching to entries in Beall's list. *ProQuest Central* provided access to the second highest number of journals from these publishers with 299 titles, or 1.41% of its holdings. In *Academic Search Complete*, 55 journal titles, or 0.40% of its holdings, were listed as being available from the six predatory publishers. Finally, in *Academic OneFile*, six journals, or 0.04% of its holdings, were available from the two predatory publishers. Table 4 identifies the quantitative title data associated with these publishers in the databases and index.

Research Question 4: In which subject areas were the predatory journals most prevalent?

As previously stated in the Methodology section, for comparison purposes, initially 14 generalized subject areas were planned to be assigned to the predatory journal titles identified in the databases and index. After subject analysis, however, no journal titles were identified in either the History or the Sports/Recreation subject area. After the subject analysis of the predatory journal titles available in the databases and index was completed, the following 12 generalized subject areas were used to represent their content:

- Arts or Entertainment
- Business
- Communication
- Education
- Government or Politics
- Humanities or Social Science

- Language or Literature
- Medicine or Health
- Multidisciplinary
- Philosophy or Religion
- Science
- Technology

Based on these generalized subject areas, the journal content associated with predatory publishers for each database and the DOAJ index were identified below. Table 5 identifies the breakdown of the journal subject coverage by total number and percentage of titles in the corresponding database or index.

Academic OneFile

Six journal titles were provided by two of the publishers on *Beall's List* in *Academic OneFile*. Based on the generalized subject categories used for this study, content from three of the titles were considered science, one was identified with medicine or health, one title with technology, and another with education. For reporting purposes, 50.0% of these journal titles were associated with science while the other 50.0% of the other titles were identified, in equal proportions, with medicine or with health, technology, and education.

Academic Search Complete

Fifty five journal titles were provided by six publishers on *Beall's List* in *Academic Search Complete*. Based on the generalized subject categories used for this study, content from 27 of the titles were related to science, 14 titles were identified with medicine or health, nine titles were associated with technology, four were multidisciplinary, and one title was related to the humanities or social sciences. For reporting purposes, 49.0% of the titles were associated with the sciences, 25.5% with medicine or health, 16.4% correlated with technology, 7.3% were identified as multidisciplinary content, and 1.8% was related to the humanities or social science.

DOAJ Index

The content from 812 journal titles was available from 123 publishers on *Beall's List* in the DOAJ index. Based on the assigned generalized subject categories, content from 258 of the titles (31.8%) were related to science, 197 titles (24.3%) were associated with technology, 197 titles (24.3%) were

TABLE 5 Distribution of predatory journals across subject categories

Subject Category	Academic OneFile		Academic Search Complete		DOAJ		ProQuest Central		Total	
	# of Titles	%	# of Titles	%	# of Titles	%	# of Titles	%	# of Titles	% of Total
Arts/Entertainment					4	0.5	1	0.3	5	0.4
Business					67	8.2	99	33.1	166	14.2
Communication					5	0.6	5	1.7	10	0.9
Education	1	16.7			26	3.2	21	7.0	48	4.1
Government/Politics					8	1.0	6	2.0	14	1.2
Humanities/Social Science			1	1.8	20	2.4	19	6.4	40	3.4
Language/Literature					1	0.1	3	1.0	4	0.3
Medicine/Health	1	16.7	14	25.5	197	24.3	43	14.4	255	21.8
Multidisciplinary			4	7.3	24	3.0	1	0.3	29	2.5
Philosophy/Religion					5	0.6	3	1.0	8	0.7
Science	3	50.0	27	49.0	258	31.8	70	23.4	358	30.5
Technology	1	16.7	9	16.4	197	24.3	28	9.4	235	20.1
Total	6	100	55	100	812	100	299	100	1172	100

identified with medicine or health, 67 of the titles (8.2%) were related to business, 26 (3.2%) were associated with education, 24 (3.0%) were multi-disciplinary, 20 (2.4%) correlated with the humanities or social sciences, eight (1.0%) were related to government or politics, five (0.6%) were communication related, five (0.6%) were associated with philosophy or religion, four (0.5%) were identified with the arts or entertainment, and one (0.1%) was related to language or literature.

ProQuest Central

Two hundred and ninety-nine journal titles were provided by 41 publishers on *Beall's List* in *ProQuest Central*. Based on the assigned generalized subject categories, content from 99 of the titles (33.1%) were related to business, 70 titles (23.4%) were associated with science, 43 titles (14.4%) were identified with medicine or health, 28 of the titles (9.4%) were related to technology, 21 (7.0%) were associated with education, 19 (6.4%) were identified with the humanities or social sciences, six (2.0%) correlated with government or politics, five (1.7%) were communication related, three (1.0%) were identified with language or literature, three (1.0%) correlated with philosophy or religion, one (0.3%) was associated with the arts or entertainment, and one (0.3%) was identified as multidisciplinary.

DISCUSSION

At the onset of this study, we did not know to which extent predatory journals existed in the three library databases and the directory/index we planned to examine. Our final data show that our worst fears were unfounded; while not wholly negligible, the three journal databases included in this study could hardly be said to be inundated with predatory journal content (although the proportion found in the DOAJ index should be of concern). The overall percentage of publishers on *Beall's List* among the total number of publishers in the databases and index (0.05–2.25, see Table 2), as well as the percentage of journals among the total number of journals in the library products we examined (0.04–8.36, see Table 4), can be regarded as relatively low.

A break-down of the findings by databases/index, and by subject areas, reveals a picture that gives more reason for concern. Predatory publishers are particularly prevalent among those whose journals are indexed in DOAJ. While 2.25% may seem a rather small number, these publishers tend to be prolific, as the much higher percentage (8.36%) of predatory journals among the DOAJ journals indicates. Assuming roughly the same number of articles per volume in both predatory and non-predatory journals, about every

twelfth recent article citation pulled from DOAJ may refer to questionable content!

The break-down by subjects shows that there is clearly more of a presence of predatory journals within certain areas: 30.5% of these journals included in our study pertain to the sciences, followed by medicine/health (21.8%), technology (20.1%), and business (14.2%). (See Table 5).

Since Beall started tracking predatory publishers and their journals, their number has grown steadily from year to year. If this trend continues, the presence of predatory content is likely to not only increase in those areas where it is a concern, but may also spread to new areas.

What does all this mean for practicing librarians—particular those in reference, instruction, collection development, interlibrary loan, and electronic resource management?

College instructors and librarians have taught students to presume that the "peer-reviewed" label stands for quality for the most part. In some cases, peer-reviewed articles have been found to be fundamentally flawed, but these were considered isolated failures within an otherwise effective quality-control process. With the emergence of predatory journals, this presumption is losing some ground. Both reference and instruction librarians need to be aware that certain subject areas and certain databases or index products may contain these journals and that the peer-reviewed label cannot always be trusted. Both need to convey to users, particularly those searching in science, medicine, health, technology, and business subjects, that predatory content may exist and that the "peer-reviewed" label should not make them assume that they are always dealing with quality content. Librarians teaching information literacy need to address quality differences among the peer-reviewed literature, including the quality problems associated with predatory journals and how to identify these journals. Especially when addressing upper-level students, library instructors might want to spend more time explaining how scholarly publishing works and sometimes failed to achieve the right level of quality, pursuing the "scholarly publishing literacy" advocated by Beall and others.[42] Library instructors may also consider workshop-type sessions addressing the campus community at large. This may prevent faculty from publishing in journals of questionable merit, and inspire a review of existing tenure and promotion policies.

Collection development and serials librarians also need to be aware of the problem posed by predatory publishing and should adjust collection policies accordingly. The absence of predatory journal content should be included among the criteria considered when deciding whether to subscribe to a particular database. Database vendors of course need to be educated on this issue, and they should be asked about their company's policy with regard to filtering out these journals in their products. Do they review the information

provided by publishers? Do they attempt to identify predatory publications, and if so, by what means? During contract (re)negotiations, librarians should keep pressure on vendors to exclude these journals from their article databases and subscription packages or negotiate prices contingent on the percentage of predatory content.

For librarians who manage electronic journal resources, this predatory problem also poses an issue. These librarians may need to be more proactive about managing link resolvers and deactivating journals from publishers known to be clearly predatory. While a thorough check for these titles puts more work on the librarian's plate, the long-term benefits of ensuring that predatory journals are not presented to end users could thus be realized. Librarians may want to consider deactivating providers of predatory content in their discovery service resource list. This also presents an issue for interlibrary loan; if a predatory article is discovered in a database search that is not full-text, the article may be requested through the interlibrary loan service, so that the library may end up paying for a "questionable" article to both the database vendor and the lending entity, with some of the proceeds going to the disreputable publisher. To avoid this, librarians could configure their interlibrary loan client to route article requests for predatory journal content to queues specifically set up to manage these requests.

LIMITATIONS

We are aware of the limitations of our study and that the data collected shows only a snapshot in time—May 2014 and September 2014. Even though this snapshot indicates that although the overall portion of predatory publisher content in the examined products are low, the potential adverse effect on libraries and end users will only grow if left unchecked. Our literature review findings indicate that predatory journals are likely to increase in number over time.

CONCLUSION

This study looked at the presence of predatory journals in library databases and attempted to determine if those journals put students at risk of including questionable content in their research papers and assignments. Our final data illustrate that the break-down by certain subjects shows there is clearly more of a presence of predatory journals within the subject areas of science, medicine/health, technology, and business in the three databases that were examined—EBSCO's *Academic Search Complete*, Proquest's *Central*, and Gale's *Academic One File*—and the proportion was found to be even higher

in the DOAJ index. Despite the limitations of this study, we believe we have uncovered a significant presence of predatory content in library subscription databases, which may already have had a deleterious effect on student research. The results may also serve as a call to librarians to be vigilant in identifying predatory content as it relates to library instruction, collection development, electronic journal management, interlibrary loan, and vendor negotiations. Future research that looks at the impact predatory publishers have on student research and information literacy would provide a stronger knowledge base for library/information literacy instruction. Studies like this should be conducted periodically to monitor the spread of predatory journals.

NOTES

1. Jeffrey Beall, "Predatory Publishers are Corrupting Open Access," *Nature* 489, (2012b): 179. doi:10.1038/489179a.

2. Budapest Open Access Initiative, 2002, http://www.budapestopenaccessinitiative.org/read (accessed May 2, 2015).

3. Bethesda Statement on Open Access Publishing, 2003, http://legacy.earlham.edu/~peters/fos/bethesda.htm (accessed May 2, 2015).

4. Ibid.

5. Higher Education Funding Council for England (HEFCE), "Policy for Open Access in the Post-2014 Research Excellence Framework," http://www.hefce.ac.uk/pubs/year/2014/201407/ (accessed July 30, 2015).

6. Mary Ellen McIntire, "Librarians Leap to the Aid of Researchers Whose Funding Will Soon Depend on Open Access," *The Chronicle of Higher Education*, July 30, 2015, http://chronicle.com/article/Librarians-Leap-to-the-Aid-of/231997/ (accessed May 2, 2015).

7. Steven Totosy de Zepetnek and Joshua Jia, "Electronic Journals, Prestige, and the Economics of Academic Journal Publishing," *CLCWeb: Comparative Literature and Culture* 16, no. 1 (2014): 1–13; E. C. Moylan, S. Harold, C. O'Neill, and M. K. Kowalczuk, "Open, Single-blind, Double-blind: Which Peer Review Process Do You Prefer? Which Peer Review Process Do You Prefer?" *BMC Pharmacology and Toxicology* 15 (2014): 55–60. doi:10.1186/2050-6511-15-55; Catherine Sharp, "Building a Successful Service: Developing Open Access Funding and Advocacy at University College London," *The Serials Librarian* 67 (2014): 276–288. doi:10.1080/0361526X.2014.954298; Abby Colbridge, "Open Access: Progress, Possibilities, and the Changing Scholarly Communications Ecosystem," *Online Searcher* 38, no. 2 (2014): 42–52; Bernard Forgues and Sebastien Liarte, "Academic Publishing: Past and Future," *M@n@gement* 16, no. 5 (2013): 739–756. doi:10.3917/mana.165.0739. Matt Enis, "Uncommonly Open," *Library Journal* 138, no. 11 (2013): 38–40. Stevan Harnad, "The Self-Archiving Initiative: Freeing the Refereed Research Literature Online," *Nature* 410, (2001): 1024–1025. doi:10.1038/35074210; M. M. Quinn, "Open Access in Scholarly Publishing: Embracing Principles and Avoiding Pitfalls," *The Serials Librarian* 69, no. 1 (2015): 58–69. doi:10.1080/0361526X.2015.1036197.

8. Jeffrey Beall, "Unethical Practices in Scholarly, Open-Access Publishing," *Journal of Information Ethics* 22, no. 1 (2013): 11–20. doi:10.3172/JIE.22.1.11; Sarah Beaubien and Max Eckard, "Addressing Faculty Publishing Concerns with Open Access Journal Quality Indicators," *Journal of Librarianship and Scholarly Communication* 2, no. 2 (2014). doi:10.7710/2162-3309; Declan Butler, "Investigating Journals: The Dark Side of Publishing," *Nature* 495 (2013): 433–435. doi:10.1038/495433a; Frank Truth, "Pay Big to Publish Fast: Academic Journal Rackets," *Journal of Critical Education and Policy Studies* 10, no. 2 (2012).

9. Adam Etkin, "A New Method and Metric to Evaluate the Peer Review Process of Scholarly Journals," *Publishing Research Quarterly* 30, no. 1 (2014): 23–38. doi:10.1007/s12109-013-9339-y.

10. Declan Butler, "Investigating Journals: The Dark Side of Publishing," *Nature* 495 (2013): 433–435. doi:10.1038/495433a.

11. Etkin, *A New Method*, 23–38.

12. Ibid., 23–38.

13. Stevan Harnad, "The Self-Archiving Initiative: Freeing the Refereed Research Literature Online," *Nature* 410, (2001): 1024–1025. doi:10.1038/35074210.

14. Zepetnek et al., *Electronic Journals, Prestige*, 7.

15. Thomas P. Gould, *Do We Still Need Peer Review?* (Lanham, MD: Scarecrow Press, 2013).

16. Ana Fresco-Santalla and T. Hernández-Pérez, "Current and Evolving Models of Peer Review," *The Serials Librarian* 67 (2014): 373–398. doi:10.1080/0361526X.2014.985415.

17. Michael Q. Dudley, "The Curious Case of Academic Publishing," *Partnership: The Canadian Journal of Library and Information Practice and Research* 8, no. 1 (2013).

18. A. O. Omobowale, O. Akanle, A. I. Adeniran, and K. Adegboyega, "Peripheral Scholarship and the Context of Foreign Paid Publishing in Nigeria," *Current Sociology* 62, no. 5 (2014): 666–684. doi:10.1177/0011392113508127.

19. Ibid., 11.

20. Scott Warren and Kim Duckett, "Why Does Google Scholar Sometimes Ask for Money? Engaging Science Students in Scholarly Communication and the Economics of Information," *Journal of Library Administration* 50, (2010): 349–372. doi:10.1080/01930821003667021.

21. Warren et al., *Why Does Google Scholar*, 349–372.

22. Warren et al., *Why Does Google Scholar*, 439–372.

23. Margy MacMillan and Allison MacKenzie, "Strategies for Integrating Information Literacy and Academic Literacy: Helping Undergraduate Students Make the Most of Scholarly Articles," *Library Management* 33.8(9): 525–535. doi:10.1108/01435121211279885.

24. Linlin Zhao, "Riding the Wave of Open Access: Providing Library Research Support for Scholarly Publishing Literacy," *Australian Academic & Research Libraries* 45, no. 1 (2014): 3–18. doi:10.1080/00048623.2014.882873.

25. Ibid., 3–18.

26. Michael Q. Dudley, "The Curious Case of Academic Publishing," *Partnership: The Canadian Journal of Library and Information Practice and Research* 8, no. 1 (2013).

27. Monica Berger and Jill Cirasella, "Beyond Beall's List: Better Understanding Predatory Publishers," *College & Research Libraries News* 76, no. 3 (2015): 132–135.

28. Jeffrey Beall, "List of Publishers: Beall's List," *Scholarly Open Access*, 2015, http://scholarlyoa.com/publishers/ (accessed March 17 and April 21, 2015).

29. Erin Bornemann, "Exposing Predatory Publishers," *Information Today* 30, no. 6 (2013): 13.

30. Ibid.

31. Jeffrey Beall, "Five Scholarly Open Access Publishers," *The Charleston Advisor* 13 (2012): 5–10. doi:10.5260/chara.13.4.5.

32. Bornemann, "Exposing Predatory Publishers," 13.

33. Walt Crawford, "Ethics and Access 1: The Sad Case of Jeffrey Beall," *Cites & Insights* 14, no. 4 (2014): 1–14; Monica Berger and Jill Cirasella, "Beyond Beall's List: Better Understanding Predatory Publishers," *College & Research Libraries News* 76, no. 3 (2015): 132–135.

34. Gale Cengage Learning, "Academic OneFile," *Gale Cengage Learning*, n.d. http://www.cengage.com/search/productOverview.do?Ntt=onefile│129663692219470930731132330063644714198&N=197&Ntk=APG%7CP_EPI&Ntx=mode+matchallpartial#Overview (accessed March 16, 2015).

35. EBSCO, "Academic Search Complete," *EBSCO*, 2015, http://www.ebscohost.com/academic/academic-search-complete (accessed March 16, 2015).

36. Jeffrey Beall, "List of Publishers: Beall's List," *Scholarly Open Access*, 2015, http://scholarlyoa.com/publishers/ (accessed March 17 and 21 April 2015).

37. DOAJ, "Directory of Open Access Journals," *Directory of Open Access Journals*, 2015, http://doaj.org/ (accessed March 16, 2015).

38. Ibid.

39. ProQuest, "Databases," *ProQuest Central*, n.d., http://www.proquest.com/libraries/academic/databases/ProQuest_Central.html (accessed March 16, 2015).

40. Jeffrey Beall, "List of Publishers: Beall's List," *Scholarly Open Access*, 2014, http://scholarlyoa.com/publishers/ (accessed May 19, 2014).

41. EBSCO, "Title Lists," EBSCO, 2014, http://www.ebscohost.com/title-lists (accessed May 19, 2014); Gale Cengage Learning, "Database Title Lists," *Gale Cengage Learning*, 2014, http://solutions.cengage.com/Gale/Database-Title-Lists/ (accessed September 30, 2014); DOAJ, "Download Metadata," *Directory of Open Access Journals*, 2014, http://doaj.org/faq#metadata (accessed May 20, 2014); Kim Levicki, e-mail message to Jennifer Huffman, May 22, 2014.

42. Linlin Zhao, "Riding the Wave of Open Access," 3–18.