Summer 7-5-2022

Full-Text Retrievals and EBSCO Discovery Service: Assessing Usage of E-Journals across Multiple Platforms

Simon Robins
University of Dayton, srobins1@udayton.edu

Full-text retrievals and EBSCO Discovery Service: Assessing usage of e-journals across multiple platforms

Simon Robins, MLIS, Coordinator of Electronic Resources & Discovery

University of Dayton Libraries, Dayton, Ohio

srobins1@udayton.edu

https://orcid.org/0000-0002-3264-5968

Full-text retrievals and EBSCO Discovery Service: Assessing usage of e-journals across multiple platforms

Abstract: This study utilizes COUNTER 5 data from the University of Dayton (UD) to measure full-text retrievals of e-journal articles from five major academic journal publishers (Taylor & Francis, SAGE, Oxford, Wiley, and Springer). Usage data from these publishers' e-journals within EBSCO is compared to the same content when accessed from publisher platforms such as Wiley Online Library or SpringerLINK. Building on previous studies that have largely focused on links (or referrals) from the library discovery layer to publisher platforms, this study analyzes usage of full text-articles stored within EBSCO Discovery Service and EBSCOhost subject databases to consider how these full-text holdings within EBSCO might affect referrals to publisher platforms. The findings indicate that full-text article holdings within EBSCO are used more often than the same content in publisher platforms, suggesting that UD students and researchers rely heavily on--and likely often start with--EBSCO for their learning and research.

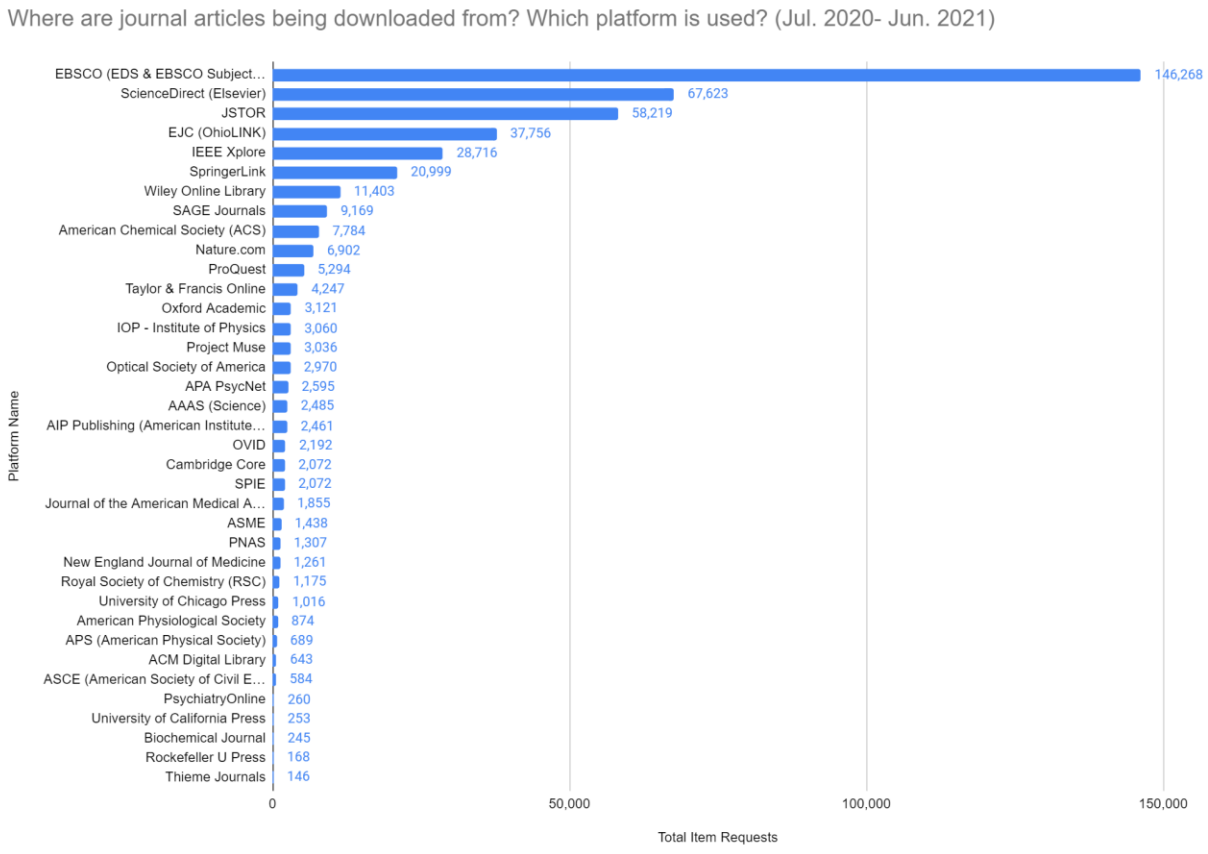
[Keywords: Discovery layers; Publisher platforms; Wiley; SAGE; Taylor and Francis; Springer; Oxford University Press; EBSCO Discovery Service; Electronic journals; Usage data; Data visualization]

Introduction

The University of Dayton Libraries (UD Libraries) is a comprehensive academic library system providing services and collections to support the university's students and teaching faculty. With a current enrollment of roughly 12,000 students, 8,200 of which are undergraduates, the university is considered a medium-sized university. In the area of library collections, UD Libraries provide vast holdings to support curricular and research needs. This comes as a result of the university being a charter member of OhioLINK—Ohio's academic library consortium and one of the largest library consortiums in the entire country—and also through its extensive collections managed independently of OhioLINK. Through consortial holdings and local subscriptions, the libraries provide access to a collection of roughly 1.2

million e-books, over 97,000 e-journals, over 40,000 streaming videos, and a print collection of roughly 560,000 titles.

Every year, UD Libraries analyzes usage of these collections via COUNTER release 5 metrics. For the past two fiscal years (running June through July), across the university's entire collection of scholarly academic e-journals, EBSCO's full-text journal holdings[1] have recorded the highest number of full-text article retrievals, beating out other platforms by significant margins (see Figure 1). Usage of EBSCO has been so high that even the second and third most-used platforms for downloading journal articles—ScienceDirect and JSTOR—did not surpass EBSCO's usage, even when combined.



[1] Within this study, any mention of EBSCO's "full-text holdings" refers to the full-text articles stored within EBSCO's subject databases and EBSCO Discovery Service (EDS), and it does not refer to custom links within EBSCO search results.

Figure 1 - Usage across all major academic e-journal subscriptions at the University of Dayton

This trend is especially surprising considering the recent (and ongoing) discussions within the e-resources community concerning how discovery layers and vendor platforms are being used. As has been argued by several prominent voices in the field (Cummings, 2021; Dempsey, 2020; Hayman, 2017; Lean Library, 2021; Nicholas et al., 2017; Robinson et al., 2013)—as well as by OhioLINK in their whitepaper (Evans & Schonfeld, 2020)—usage of EDS and of EBSCO databases such as CINAHL and Academic Search Complete typically constitute one platform, alongside many other platforms, that faculty and students use to access literature for their learning and research. Importantly, these authors assert that usage of library interfaces are unlikely to be the most prominent search platforms used. To back this point up, many of these studies have relied on referral data from vendor platforms such as ScienceDirect, JSTOR, or IEEE to show how relatively small numbers of researchers are using discovery layers or other library-controlled subject databases to reach these vendor platforms (Cummings, 2021; Evans & Schonfeld, 2020; Lean Library, 2021). The resulting argument is that students and faculty might not be using these spaces as their starting point when searching for their projects and assignments, and they are more likely reaching the full-text held within the publisher platforms via other search interfaces, most often Google and Google Scholar (Cummings, 2021; Evans & Schonfeld, 2020; Hayman, 2017; Nicholas et al., 2017; Robinson et al., 2013).

Keeping all of this in mind, a decision was made to delve deeper to examine what this EBSCO usage at UD actually represents in terms of the publishers and e-journals being accessed. A further analysis was also conducted to see how this usage relates to retrievals of the same e-journal content within vendor-controlled publisher platforms that UD subscribes to. This ultimately led to an in-depth analysis of usage trends from July 1st, 2020 to June 30th, 2021

within 5 major publishers—the most highly used within EBSCO's full-text holdings: Taylor & Francis, Wiley, SAGE, Oxford, and Springer. Usage of e-journal titles from these five publishers were compared across the different access points available to UD students, faculty and staff; namely, EBSCO's full-text holdings, the vendor platforms, the OhioLINK Electronic Journal Center's (EJC) full-text holdings, and JSTOR.

This study will explain the methodology used to gather and clean this usage data, as well as the methodology used to create a series of visualizations illustrating trends found within the data. There will also be a discussion of what these usage trends suggest about student and faculty user behavior. Lastly, and perhaps most importantly, this study will explore a potential gap in existing research and call for further studies to be conducted at other universities to assess whether or not the patterns seen here are an anomaly. If the usage discussed here is occurring more widely, then it could reveal a significantly different type of usage pattern than what has been previously discussed, one where the end user is more likely to start in the discovery layer or subject database and directly download full-text articles from within library-managed search interfaces and less likely to link out from these interfaces to external publisher platforms.

Literature Review

The majority of studies assessing where students and faculty are conducting their online searches to access e-journals have either relied on vendor-supplied analytics data or on survey data from researchers and students (Cummings, 2021; Dempsey, 2020; Evans & Schonfeld, 2020; Hayman, 2017; Lean Library, 2021; Nicholas et al., 2017; Robinson et al., 2013). Both forms of data have typically indicated relatively low numbers of users accessing or preferring library-managed databases or discovery layers as a means to discover full-text

journal articles. These trends have prompted some to question the long-term utility of discovery layers. For example, Joel Cummings (2021) at Washington State University discussed how at his institution only 10% of all referral data is coming directly from library discovery layers when users are accessing publisher platforms to access and download full-text (p. 69). As they point out, the vast majority of referral traffic to these anonymous publisher platforms and subscriptions[2] is either a direct referral or coming from Google or Google Scholar (p. 69). The article goes on to suggest that these "findings may raise questions about the traditional and significant investments libraries make in these [discovery] resources" (p. 72) and asks the question: "if libraries and/or library vendors may be able to develop more or fewer or improved tools, what must these tools be to provide to users of the research journal literature a functionality not found via Google and Google Scholar?" (pp. 72–73).

If there's one repeated trend across Cummings' study and the additional studies and whitepapers mentioned in the previous paragraph, it's the way the authors focus on the means by which users are reaching the publisher platforms containing the full text. Another way of phrasing this is to ask where the users are "starting" their searches to reach the full text held within the publisher platforms, to see if users are relying on commercial search engines, Google Scholar, social media, or library resources such as databases and discovery layers. Several of these studies have reported on survey findings where researchers and students have indicated that they prefer commercial search engines, Google Scholar, or other resources such as PubMed as their starting point for research and assignments (Inger & Gardner, 2016; Lean Library, 2021). These surveys indicate how researchers and students are increasingly not using the library databases at all.

[2] The e-journal publishers and vendors analyzed in this study are anonymous because the publishers required anonymity as a condition of sharing the data.

In two other studies, usage of search interfaces is assessed more broadly by looking at link resolver and interlibrary loan data, with the authors considering the impact that the discovery layer has on usage over time (Wang, 2020; Wang & Howard, 2012). Again, these studies focus on the number of link outs occurring from search platforms to external sources. At the same time, in both cases the authors do take a slightly different route by utilizing link-out data pulled directly from their institution's custom link resolver, as well as data from their interlibrary loan referral traffic, rather than relying on web analytics data provided by the publishers. Using this data allowed them to more fully assess whether this ILL or link resolver traffic was coming from their discovery layer, from the EBSCO database Academic Search Complete, or from Google Scholar, and it also allowed the authors to show how much the discovery layer and Academic Search Complete were used, demonstrating that the link resolver and ILL received more referrals from these library-managed products than they did from Google Scholar. Yet, at the same time, unlike the studies discussed in the previous paragraphs, using this data likely excluded certain users who reach publisher platforms through IP range authentication when on campus, passing by the library systems entirely. Moreover, while these two studies assessing internally-provided usage data provide a different angle, they still do not consider this usage of custom links alongside the number of full-text downloads occurring within Academic Search Complete or the Discovery Layer, nor do they incorporate any discussion of the usage rates occurring within publisher platforms such as Wiley Online Library.

In a slightly separate area, there are several studies measuring the overall impact of a discovery layer implementation, specifically considering the impact it has had on e-resource usage (Calvert, 2015; Evelhoch, 2016; Ngo et al., 2019). All of these studies examine the download rate of articles, alongside other factors such as the impact of a discovery layer on print

circulation, while focusing less on whether users are linking from the discovery layer to the external publisher platforms. In the case of Calvert, the author measures the download rate of full-text content within EBSCO, as well as the impact the discovery layer implementation had on the download-rate within specific journal platforms such as ScienceDirect, Sage, and Wiley. They do this by presenting the download rates within both areas. And while Calvert does present this data, they do not consider why certain packages such as Wiley, Project Muse, and the American Chemical Society saw significant decreases in the number of full-text downloads occurring within their respective publisher platforms in the year after implementing EBSCO EDS (p. 91). In terms of the impact that these discovery layer implementations had on overall usage of e-journal platforms, the results were mixed: Evelhoch reported a decline in journal article requests at Central Washington University after implementing Primo, Calvert reported a sharp increase in e-journal requests across the majority of publisher platforms at Western Carolina University after implementing EDS, and Ngo et al. found that more publisher platforms saw a decrease in the number of article requests than those seeing an increase.

Two studies assessing the impact of discovery layers on e-journal usage within publisher platforms conclude their publications by suggesting future research to address a gap (Levine-Clark et al., 2014; Ngo et al., 2019). Specifically, the gap suggested by both studies points toward a need to combine the concerns raised by studies examining referral data (studies confirming that referrals are low) with the concerns raised by studies examining the overall impact of discovery layers on e-journal usage (that more often than not, discovery layer implementations can have a noticeable effect—positive or negative—on e-journal downloads within publisher platforms). In their longitudinal study examining the impact of discovery layer on e-journal usage within publisher platforms, Levine-Clark et al. (2014) conclude by noting

how "future phases of [their] study will incorporate aggregator-hosted journal content" alongside their earlier analysis of the discovery layer's impact on usage of publisher-hosted journal content (Levine-Clark et al., 2014, p. 256). Similar to the trends noted in the previous paragraph, their study found that any increases in usage within e-journal publisher platforms varied, depending on the different discovery layer implemented (Summon, Primo, EDS, or WorldCat Local), but they also conclude that based on their data the discovery layer was undoubtedly impacting usage (without determining why the impact was occurring).[3] Worth noting, this study analyzed usage across multiple libraries in several different countries and across multiple different platforms.[4]

Ngo et al. (2019) was more concrete when they suggested in their conclusion about the impact of EDS at UC Berkeley that "full-text availability in the EDS interface may in fact negatively correlate with usage reported by a publisher or platform" and called for future studies to assess this assumption (Ngo et al., 2019, p. 236). They also go on to point out how Marshall Breeding questioned whether discovery layers are ranking their full-text content higher than other results in their relevancy ranking, increasing the likelihood that users will download these PDFs, rather than linking to the publisher platforms (Breeding, 2015). The authors agree with Breeding, but question whether there will ever be transparency from companies like EBSCO and ProQuest to more fully understand their relevancy rankings. The conclusions and calls for further research expressed by both studies point to an important question: if the implementation of a discovery layer has a noticeable impact on usage of e-resources outside of the discovery layer, then how do full-text holdings provided directly within the discovery layer correlate to these declines and increases within individual publisher platforms? As of December 2021, neither of

[3] The authors of this study explicitly state that they did not seek to find out why the discovery layer was impacting usage, but rather sought to determine whether it impacted usage.
[4] "Twenty-eight libraries from the US were included, two from Canada, and one each from the UK, Australia and New Zealand" (Levine-Clark et al., 2014, p.250)

these two earlier studies have published any sort of follow-up, nor does there seem to be any additional authors exploring this gap. This study seeks to take the next step toward filling the gap by assessing usage at one medium-sized, research university.

Materials and Methods

Measuring Article Retrievals within EDS and the broader "EBSCO ecosystem"

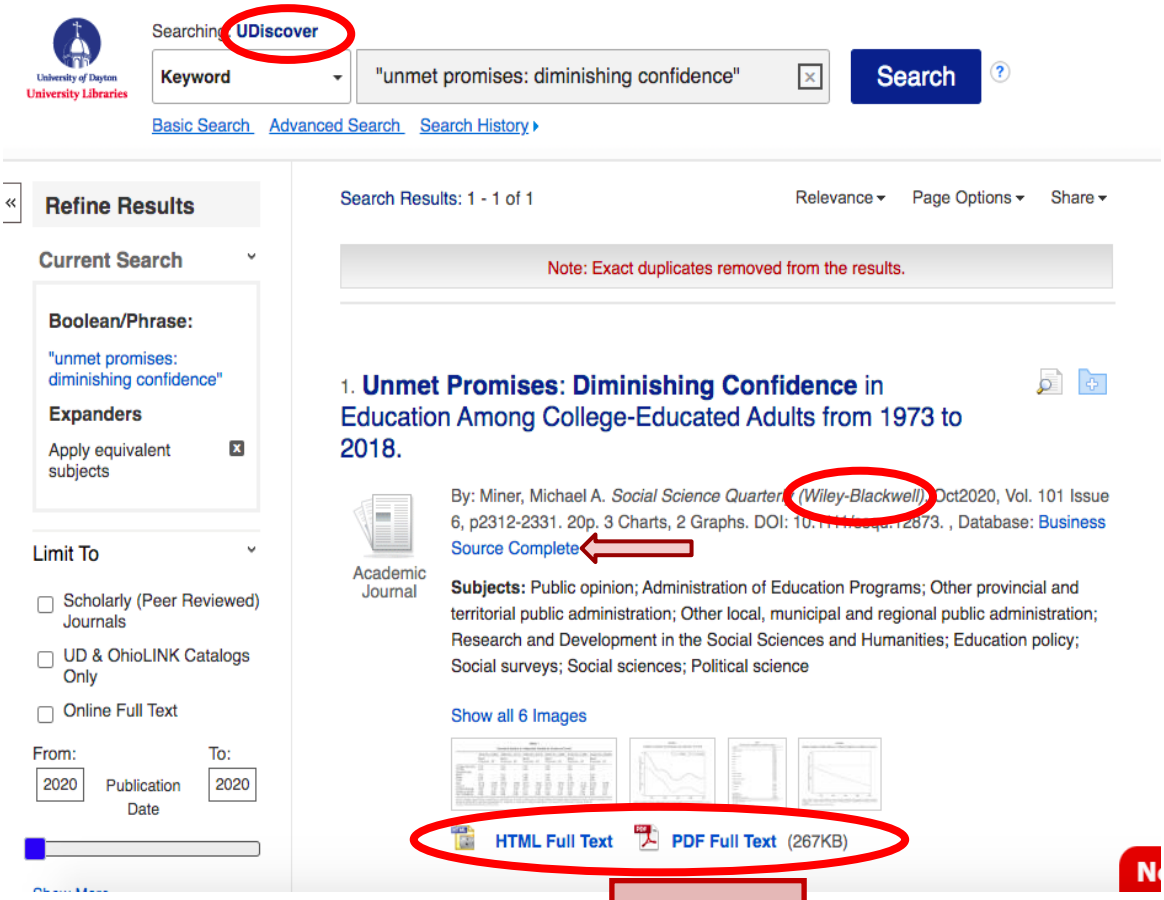
Within this study, any reference to full-text downloads (or retrievals) occurring within EBSCO's full text holdings refers to the direct downloads of PDFs (or the viewing of the full text in the browser) directly from within EBSCO's interface. For example, retrievals are being measured when a user downloads the PDF of a journal article from EBSCO's Business Source Complete database (see Figure 2). Importantly, the usage represented in the previous section (displayed in Figure 1) and in all subsequent sections does not measure "CustomLinks" (how EBSCO labels referrals) from EBSCO's databases or UD's iteration of EDS, UDiscover, to external publisher platforms.

At the same time, any measure of EBSCO downloads does account for EBSCO's ability to feed the full-text from one EBSCO database into the other. This occurs because EBSCO's extensive collection of subject databases and its EDS constitute a highly interconnected set of indexes and full-text databases that create a broader network, providing the end user with a situation where if the full-text isn't available within one specific EBSCO subject database, then it might be available in another. On top of this, the full-text holdings from specific EBSCO subject databases such as Business Source Complete will directly appear within the search results of EDS, creating an experience where the end-user can directly download the full-text from

Business Source Complete (or any other of EBSCO's vast collection of subject databases) without having to leave the EDS interface (see Figure 2).

Moreover, the vast majority of EBSCO subject databases can be customized to look indistinguishable from each other and from EDS. The only noticeable difference is the name label of the database above the search box (see Figure 2). This creates an experience where the end-user becomes highly familiar with how to use multiple EBSCO databases and float between the different search interfaces. Based on this seamless quality—and based on the fact that it is not possible to distinguish in the COUNTER 5 data whether the end user downloaded an article provided by EBSCO from within a subject database or from within EDS—both EBSCO resources will be analyzed as a single resource (or "ecosystem") within this study.[5]

[5] At the University of Dayton, we subscribe to the EBSCO expanded pack of databases and to EDS. This situation with EBSCO is not unique to UD, and the vast majority of OhioLINK institutions have the same custom set up with EBSCO and EDS to match their institution's branding and to link to their custom subscriptions.


Searching UDiscover
Keyword
"unmet promises: diminishing confidence"
Search
Basic Search
Advanced Search
Search History
Refine Results
Current Search
Boolean/Phrase:
"unmet promises: diminishing confidence"
Expanders
Apply equivalent subjects
Limit To
Scholarly (Peer Reviewed) Journals
UD & OhioLINK Catalogs Only
Online Full Text
From:
2020
Publication Date
To:
2020
Search Results: 1 - 1 of 1
Relevance
Page Options
Share
Note: Exact duplicates removed from the results.
1. Unmet Promises: Diminishing Confidence in Education Among College-Educated Adults from 1973 to 2018.
By: Miner, Michael A. Social Science Quarterly (Wiley-Blackwell). Oct2020, Vol. 101 Issue 6, p2312-2331. 20p. 3 Charts, 2 Graphs. DOI: 10.1111/ssqu.12873. , Database: Business Source Complete
Academic Journal
Subjects: Public opinion; Administration of Education Programs; Other provincial and territorial public administration; Other local, municipal and regional public administration; Research and Development in the Social Sciences and Humanities; Education policy; Social surveys; Social sciences; Political science
Show all 6 Images
HTML Full Text
PDF Full Text (267KB)

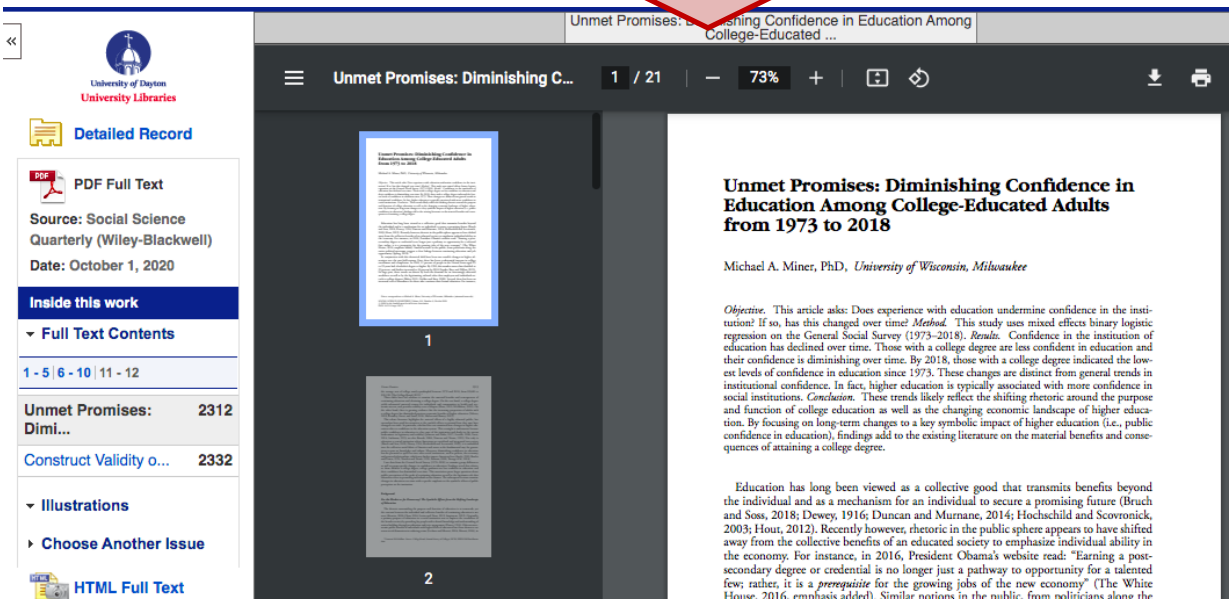

Unmet Promises: Diminishing Confidence in Education Among College-Educated ...
Unmet Promises: Diminishing C...
1 / 21
73%
Detailed Record
PDF Full Text
Source: Social Science Quarterly (Wiley-Blackwell)
Date: October 1, 2020
Inside this work
Full Text Contents
1 - 5 | 6 - 10 | 11 - 12
Unmet Promises: Dimi... 2312
Construct Validity o... 2332
Illustrations
Choose Another Issue
HTML Full Text
Unmet Promises: Diminishing Confidence in Education Among College-Educated Adults from 1973 to 2018
Michael A. Miner, PhD, University of Wisconsin, Milwaukee
Objective. This article asks: Does experience with education undermine confidence in the institution? If so, has this changed over time? Method. This study uses mixed effects binary logistic regression on the General Social Survey (1973–2018). Results. Confidence in the institution of education has declined over time. Those with a college degree are less confident in education and their confidence is diminishing over time. By 2018, those with a college degree indicated the lowest levels of confidence in education since 1973. These changes are distinct from general trends in institutional confidence. In fact, higher education is typically associated with more confidence in social institutions. Conclusion. These trends likely reflect the shifting rhetoric around the purpose and function of college education as well as the changing economic landscape of higher education. By focusing on long-term changes to a key symbolic impact of higher education (i.e., public confidence in education), findings add to the existing literature on the material benefits and consequences of attaining a college degree.
Education has long been viewed as a collective good that transmits benefits beyond the individual and as a mechanism for an individual to secure a promising future (Bruch and Soss, 2018; Dewey, 1916; Duncan and Murnane, 2014; Hochschild and Scovronick, 2003; Hout, 2012). Recently however, rhetoric in the public sphere appears to have shifted away from the collective benefits of an educated society to emphasize individual ability in the economy. For instance, in 2016, President Obama's website read: "Earning a post-secondary degree or credential is no longer just a pathway to opportunity for a talented few; rather, it is a prerequisite for the growing jobs of the new economy" (The White House, 2016, emphasis added). Similar notions in the public, from politicians alone the

Figure 2 – Top - This search result is provided through the EBSCO subject database "Business Source Complete," and it directly feeds into UD's discovery layer, UDiscover. When the user clicks on the "PDF Full Text" link, they are not taken outside of EBSCO; the full PDF is stored within EBSCO's servers and able to be viewed directly with a single click (bottom half of image)

Data gathering, data cleaning, and initial visualizations

The initial usage statistics for e-journal downloads across all UD subscriptions (seen in Figure 1) were pulled using COUNTER 5 reporting and specifically used the "Journal Requests, Excluding OA_Gold (TR_J1)" report.[6] This report was pulled for all active e-journal subscriptions at the university, either directly from the vendor platform or from Springshare's LibInsight platform. Once all COUNTER statistics had been gathered across all current UD subscriptions containing e-journals, the results were totaled and visualized within a basic excel chart. The metric of "total item requests" was used within this analysis and visual (and within all subsequent visuals). Each request (or retrieval) measures when a user downloads the PDF of a journal article, or when they view the full-text within the browser. This process established that the e-journals we subscribe to through the EBSCO Expanded Pack and EDS constitute the most highly used collection across all of our active subscriptions, and it demonstrates how these platforms are the most highly used spaces to retrieve full-text journal articles.[7] This discovery led

[6] As defined by COUNTER, a "request" is when a user accesses "the actual content item." The time period represented by the data in this visual, as well as in all subsequent visuals discussed here, is July 1st, 2020 to June 30th, 2021.

[7] This could indicate a number of different usage trends. It could mean EBSCO has far more full text holdings for e-journal titles, and therefore the wide selection of titles are resulting in a lot of retrievals. It could also indicate that our users are ending up at the full-text holdings provided by EBSCO *before* they are ending up at the vendor platforms or at the Electronic Journal Center (EJC).

to further questions and a subsequent analysis of the EBSCO TR_J1 report within excel.

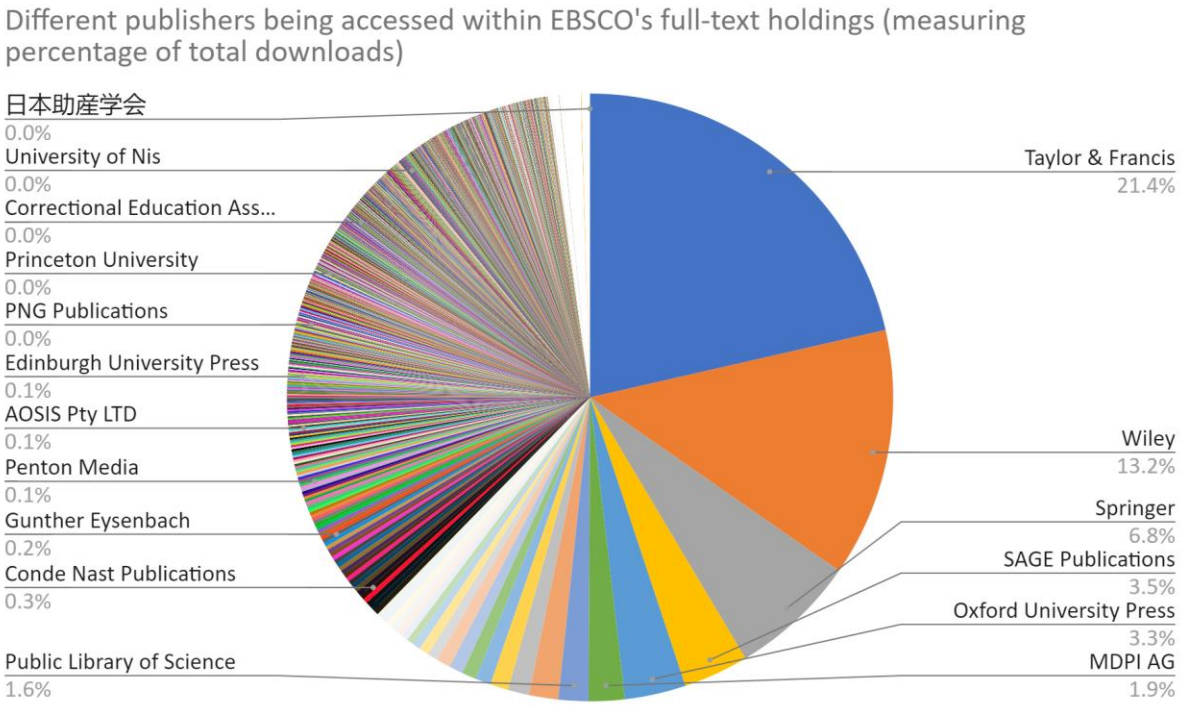


Figure 3

The logical next step was to see which publishers within EBSCO were receiving the highest number of full text retrievals. To do so, the column within the dataset labeled “publisher” was used. Once this data had been extracted, some data cleaning was required using OpenRefine to cluster the publisher titles, and Tableau and excel were used to create the visual calculating the percentages of use for each prominent publisher (Figure 3). Upon making these calculations and visualizations, it was immediately apparent that Taylor & Francis, Wiley, Springer, SAGE, and Oxford University Press were the most commonly used publishers within EBSCO’s e-journal holdings. This made any further analysis far more straightforward because UD Libraries provides access to all five of these publishers’ vendor platforms, through our local subscriptions or through our consortial subscriptions. UD students and faculty have access to at least 250

unique titles from each of the publishers.[8] These titles are in addition to any of the titles provided by EBSCO Expanded Pack or EDS, but at the same time, some of the titles and holding ranges between these two sources—EBSCO and the vendor platform—have overlap. For the end user, this last point increases the chance that a single journal article from any of these five major publishers will be available within several separate platforms: within EBSCO or within the vendor platform, while also potentially being available within OhioLINK's Electronic Journal Center (EJC), a database available to most OhioLINK members containing over 32 million full-text articles in 10,000 journals. An additional platform where UD students might have access to the same full-text journal article is JSTOR.[9] All of these factors provided a unique opportunity to see how usage compares across platforms at the journal title level, and one of the driving questions for this analysis was to see how much overlap there was between usage of individual journals across these different platforms.

Having established which five publishers to analyze, usage data was gathered from EBSCO, the EJC, each of the five publisher platforms, and from JSTOR, again using COUNTER 5's "Journal Requests, Excluding OA_Gold (TR_J1)" report for each platform and the "total item requests" metric. To accurately compare the data from the publisher platforms with the data from EBSCO, some criteria were established. First, only journals published by one of the five publishers were included; this analysis did not include publications from other publishers that

[8] UD faculty and students have access to: 2,412 subscription e-journal titles within SpringerLINK; 755 subscriptions within Taylor & Francis; 1,899 subscriptions within Wiley Online Library; 1,143 subscriptions within SAGE; and 267 subscriptions within Oxford. For inclusion in this analysis, a journal title had to be from one of the five publishers and have at least 1 full-text retrieval within the year time range. For e-journal titles accessed within one of the five vendor platforms: SpringerLINK had 1,199 titles; for Taylor & Francis, 538; for Wiley, 1,058; for SAGE, 681; and for Oxford, 250. This totals 3,726 e-journal titles accessed within the vendor platforms. Within EBSCO, 2,979 unique e-journals from these 5 publishers saw 1 or more full-text retrievals; within OhioLINK's EJC, 2,122 unique e-journals from these 5 publishers saw 1 or more full-text retrievals; and within JSTOR, 560 unique e-journals from these 5 publishers saw 1 or more full-text retrievals. Please see the appendix for a further breakdown and visualization. Across all these different platforms, there is overlap in the titles. After deduplication, we estimate roughly 5,700 unique journal titles saw at least 1 article retrieval in the timeframe examined, across all 5 publishers.

[9] Access to all of these platforms is configured with IP authentication for on-campus and with ezproxy for off-campus, and, as a result, the end user does not have to authenticate to access any of the content when on-campus or within university-owned housing.

just happen to be available within one of the vendor platforms. This decision was made so that accurate comparisons could be made between data from EBSCO, JSTOR, the EJC, and the vendor platforms. A minor exception was made to include Routledge titles within the Taylor & Francis total. This decision was made because all Routledge titles are available from within Taylor & Francis Online and from within EBSCO (and because Routledge is owned by Taylor & Francis). Conversely, even though nature.com titles are owned by Springer Nature, they were not included within the Springer analysis because none of the nature.com titles are available from the SpringerLINK platform. Within the SAGE Journals dataset, a small number of titles had to be excluded because they are not owned by SAGE. For this reason, both SAGE and SpringerLink have slightly higher numbers in Figure 1 than they do in the subsequent visuals and analysis. In the case of Wiley Online Library and Oxford Academic, the COUNTER report contained usage data solely from e-journal titles published by these two publishers. In the cases of the EJC, EBSCO, and JSTOR, subsets of the larger datasets had to be pulled in order to gather the relevant publisher's data. For all three, the large JR_1 report spreadsheet was formatted as a table so that the relevant publisher could be filtered down to only include data applying to the five major publishers.

All of this usage data was then combined into a single spreadsheet and was cleaned using OpenRefine. Specifically, the journal titles needed to be clustered; for example, the journal "New Media & Society" was clustered with "New Media and Society." Once the data was cleaned, analysis was conducted within Tableau. Tableau allowed for a more in-depth set of visualizations, by layering the various data points within a single image. The following section will present several images from this analysis within Tableau, with brief discussion of what the data suggests.

Results

Visualization of usage across the different platforms

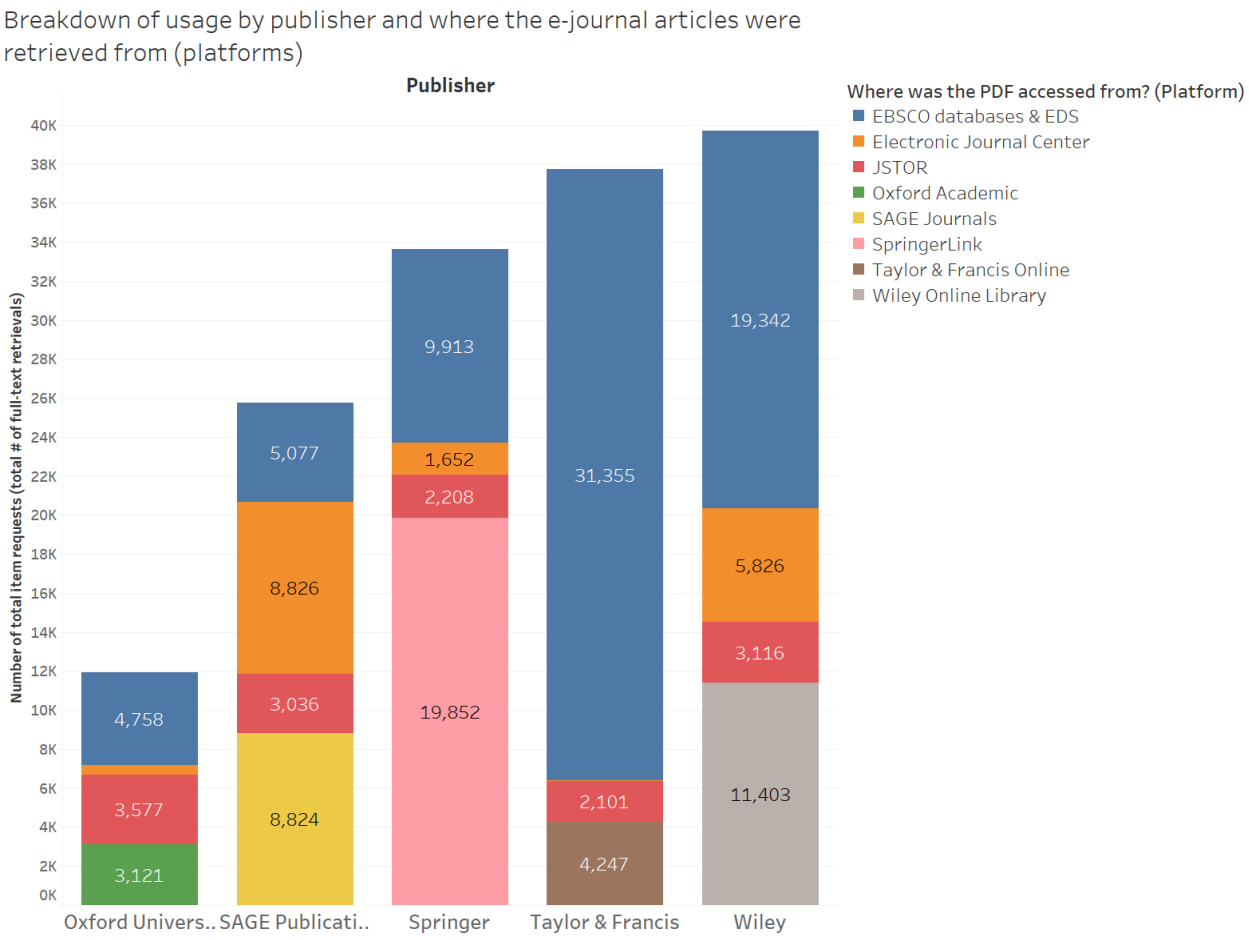


Figure 4

Figure 4 displays where full text articles from each of the five publishers were retrieved from, in terms of the platform used. The publisher with the highest number of full-text retrievals was Wiley, with 19,342 retrievals (or roughly 49% of all retrievals) occurring within EBSCOhost databases or within EDS. A further 5,826 retrievals of Wiley content (or 15% of all retrievals) occurred within the EJC. Only 11,403 retrievals (or 29%) occurred within Wiley

Online Library. In Taylor & Francis the usage of EBSCO to retrieve full text is even higher with 31,355 in total (or 83% of all retrievals). For SAGE and Oxford, the trend is different, with more usage occurring within the vendor platforms or in JSTOR (30% of all retrievals of Oxford journal articles occurred in JSTOR, while 34.25% of retrievals of SAGE content occurred in SAGE Journals). Strikingly, in the case of SAGE, 34.26% of full-text retrievals occurred in OhioLINK's EJC. This is far higher than the usage of any other of the publisher's content within the EJC. The major outlier is how SpringerLINK was used to access Springer e-journals; it is the only vendor platform where more than 35% of full-text retrievals occurred, with 59% of all retrievals (or 19,852 total retrievals) occurring within SpringerLINK.

Worth noting, when you combine the two places where users cannot reach journal articles via commercial search engines such as Google (EBSCO/EDS and the EJC) and compare this usage with a combination of the two places where users can reach journal articles via commercial search engines (JSTOR and one of the five vendor platforms), in every case except Springer, the combined totals of the EJC and EBSCO surpasses at least 40% of all full-text retrievals. For Wiley content, EBSCO and the EJC constitute 63% of all retrievals; for Taylor & Francis they constitute 83%; for SAGE they constitute 54%; and for Oxford they constitute 44%. Even for Springer, the two constitute a combined total of 34% of all retrievals.

This same data, shown in a different way—where the total usage within the platforms is measured—indicates the strong usage of the EBSCO platforms when compared to the five vendor platforms, JSTOR, and the EJC (see Figure 5). As stated earlier, the EBSCO interface at UD is nearly identical across the large selection of subject databases and the EDS, and users often bounce between them when they seek to expand or narrow their search, or when they are searching within one EBSCO databases or the EDS and are referred to a separate EBSCO

database to retrieve full-text. Based on this uniform and often seamless user experience, the author of this study does feel that it is appropriate to visualize this usage as a single column to emphasize the high amount of downloads occurring within this broader EBSCO ecosystem. The same cannot be said for any of the other platforms; all are owned by separate publishers and are not seamlessly interconnected and have different appearances and information architecture.

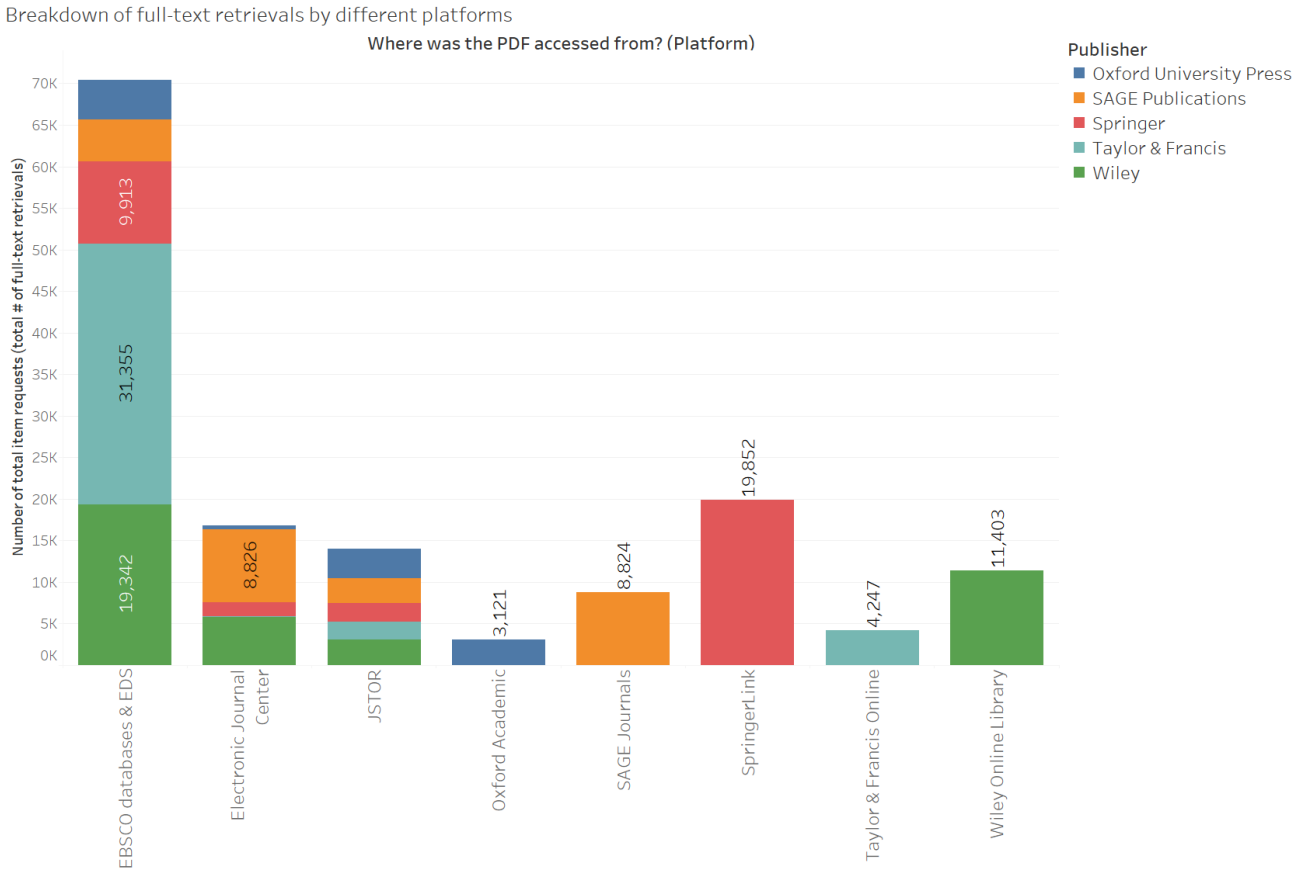


Figure 5

This last point about heavier usage within EBSCO is further indicated by Figure 6 which shows the average number of downloads (or retrievals) per journal that occurred across the various platforms. To calculate these averages, the number of downloads occurring within a specific platform for a specific publisher was divided by the number of unique journal titles accessed (those that had at least 1 article retrieval within the date range) within that platform.

Note that EBSCO and JSTOR are the two platforms where users downloaded far more articles on average per title. While this data might present more questions than it answers about user behavior, it still does indicate that certain journal titles within JSTOR and EBSCO are being accessed far more frequently than titles within the EJC or the five vendor platforms.

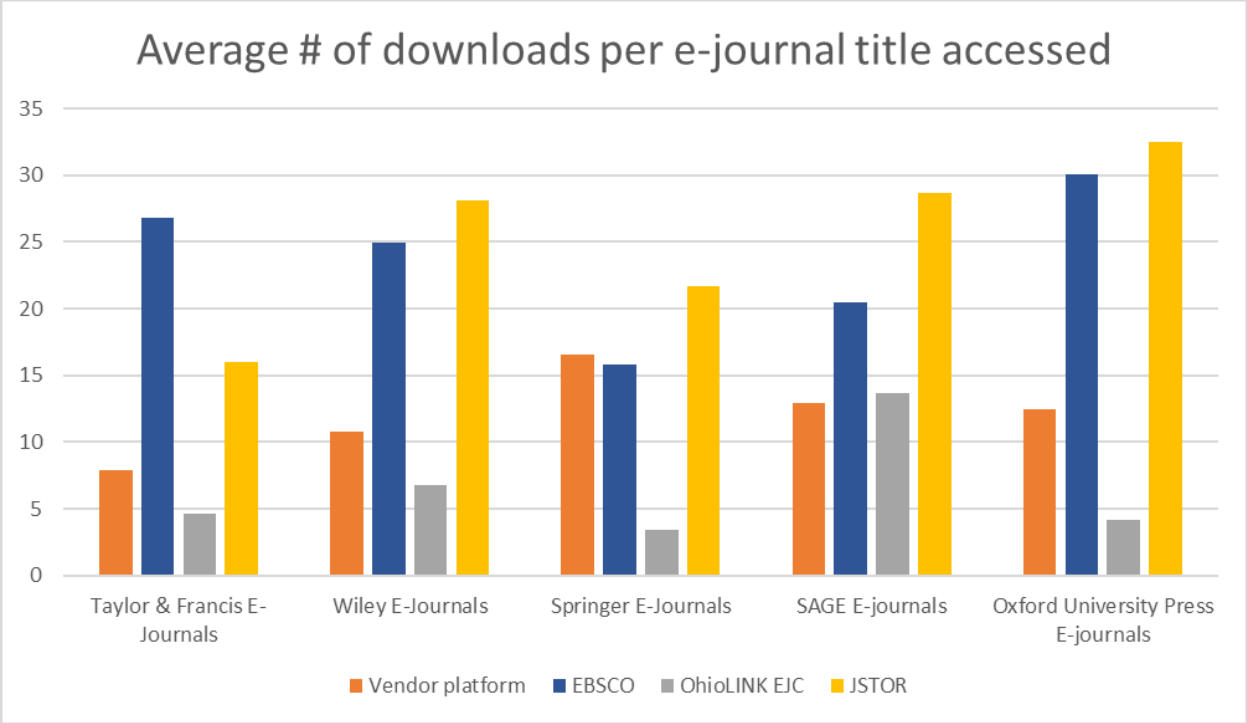


Figure 6

Visualization and Analysis at the Journal Title Level: Is there overlap between titles being accessed?

The next part of the analysis involved visualizing the usage data across all platforms within Tableau in order to see which e-journal titles were being used the most, as well as to see where the full text articles were being retrieved from in terms of the platform used. This part of the analysis proved crucial because the overall aim of the study is to verify whether or not the same titles were being used across the different platforms. Figure 7 shows all titles receiving 300

or more retrievals. Within each bar (representing the number of full-text retrievals for an individual title) usage of the title within the different platforms is indicated by different colors. Unsurprisingly, EBSCO was found to be the most prominent platform (indicated in dark blue in the visual). Across all 33 journal titles that fit the criteria for this visual, 17,133 total retrievals occurred (or 11% of all full text retrievals discussed in this study). The percentage of these full-text retrievals occurring within EBSCO was 50% (or 8,624 retrievals). SpringerLink was the second most used platform with 14% of retrievals, and the EJC had 13%.

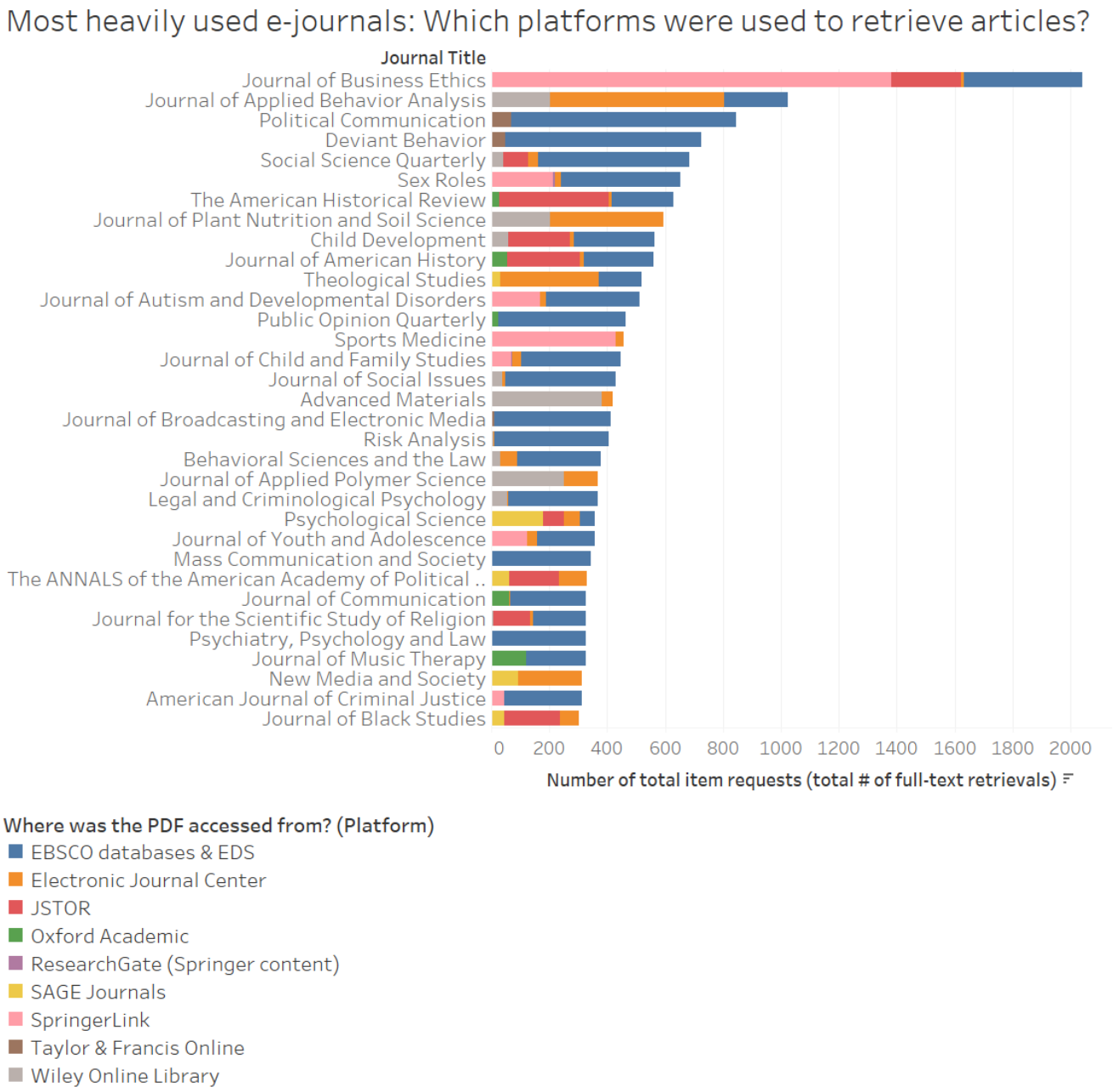


Figure 7

To make sure that this pattern within individual journal titles wasn't simply occurring within the most highly used titles, an additional visual was created displaying the platforms used across all 5,700 unique journal titles that saw at least 1 article retrieval in the timeframe examined, excluding those 33 titles visualized in Figure 7. Within this visualization (Figure 8)

EBSCO is clearly identified by the dark blue color label, displaying how widespread this pattern of users going to EBSCO to retrieve full-text was across all 5 publishers and the majority of individual titles.

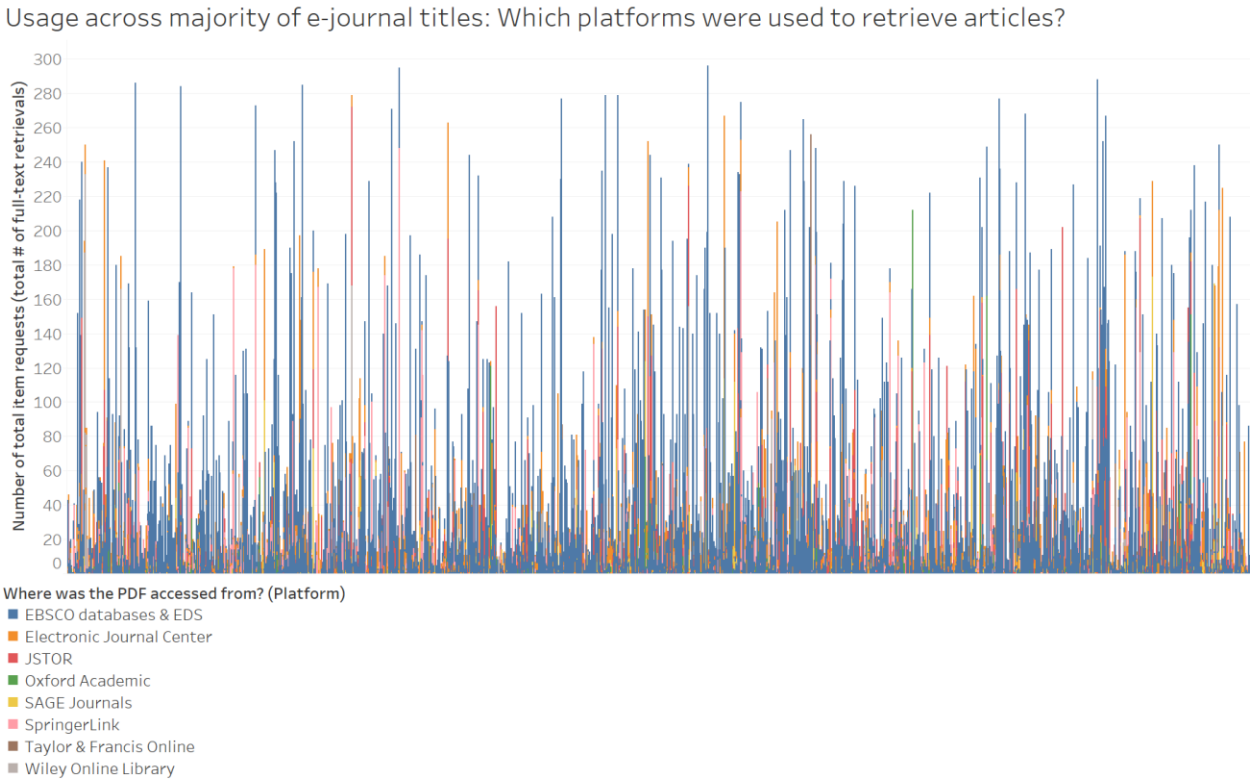


Figure 8 - Visualization of usage across all journals (across all five publishers), excluding the 33 titles visualized in Figure 7.

Having established that usage of individual journal titles within EBSCO was occurring across multiple titles, a further analysis of the overall usage was conducted to see what percentage of the total number of retrievals occurred in EBSCO (displayed in Figure 9). Across all of the platforms used to access articles from the five publishers, a combined 148,746 total item requests were recorded. Out of the 148,746, 70,445 (or 47.4%) were retrievals occurring directly within EBSCO (either a subject database or EDS). This is roughly the same percentage measured across the top 33 titles discussed above. On top of this, 16,816 (or 11.3%) occurred

within OhioLINK's EJC. Both of these spaces are not indexed by commercial search engines like Google, and the only way to reach them is either directly through the libraries' website or by authenticating into another database or platform that uses EBSCO's Full Text Finder feature. In sum, these two spaces represent a mode of access that typically requires a demonstration or tutorial on how to access and use them to access full text; it is different from being able to find a journal article indexed by JSTOR or Wiley Online Library in a set of Google search results.

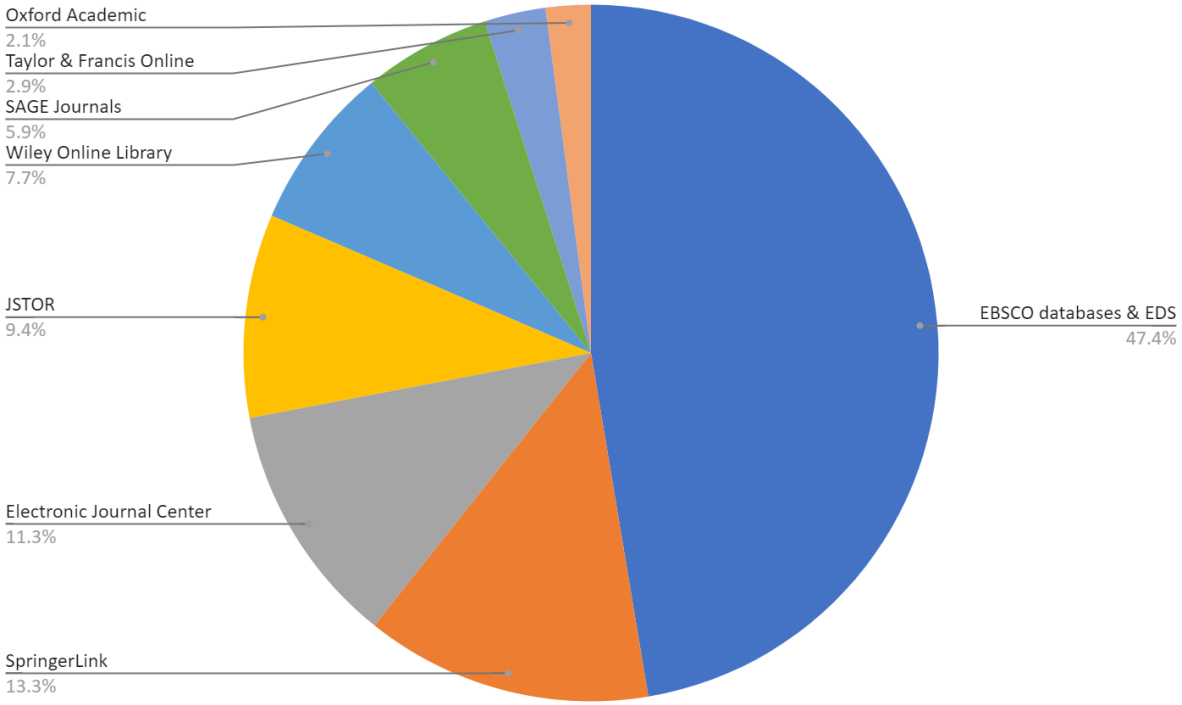


Figure 9

If we only look at the five publisher platforms, out of the total 148,746 item requests, 47,447 occurred within these platforms (or 31.8%). The remaining usage (14,038 item requests) occurred within JSTOR, which contains content from multiple publishers across many different subjects. Worth noting, based on trends reported by previous studies (Cummings, 2021), it's safe to assume that at the very least 10% of these item requests within the five publisher platforms

and in JSTOR (10% of 61,485 total item requests) likely came directly from EBSCO's custom links, either directly from EBSCO subject databases or EDS, or through Full Text Finder being linked to from these databases or other databases such as PubMed or Web of Science. Accounting for these custom links, we can estimate that roughly 62.8% of all these 148,746 item requests occurred either directly within the EJC, within an EBSCO database or EDS, or they occurred as a result of a link from EBSCO to one of the five publisher platforms.

--

Discussion & Conclusion: Why go to Wiley Online Library when you can download the article before leaving EBSCO?

This analysis strongly suggests that UD users prefer to use EBSCO over the publisher platforms, in cases where full-text journal articles are available directly from EBSCO's e-journal packages. These trends run counter to what many prominent voices within the field of e-resources are suggesting, and they likely come as a result of this paper's focus on a gap in previous studies: while many studies have considered whether or not discovery layers or subject databases are referring users to publisher platforms (through custom links and link resolvers), none have fully considered the impact that duplicate full-text holdings have on referral traffic. If a duplicate full-text PDF is immediately available from EBSCO's servers, then why would the end user go to a publisher platform like Wiley Online Library or Taylor & Francis Online to retrieve it? And wouldn't this trend cause the referrals from EBSCO to platforms like Wiley Online Library to be low?

This last point is contingent on users choosing to start their searches within the EBSCO databases or EDS, and the data discussed here strongly suggests this trend at UD. To further

support these conclusions, it is worth noting that the vast majority of the content from these five publishers when provided as full-text within EBSCO is embargoed (typically 1–2 years), while the same content when provided within the five publisher platforms typically has no embargo.[10] On top of this, EBSCO's journal article records are not indexed by Google or other search engines, and the only way to reach the EBSCO-housed PDFs mentioned in this study is through the library's website or from a direct permalink (or through EBSCO's Full Text Finder).

It is also worth noting that in no way are the trends discussed in this paper an aberration. Usage has surprisingly remained consistent within EBSCO, even with the COVID-19 pandemic shutdowns: in calendar year 2018 we recorded 153,523 total item requests, in calendar year 2019 161,548 total item requests, and in calendar year 2020 there were 156,018 total item requests. In fall 2021, UD's students, faculty, and staff fully returned to campus, while in the spring of that year, the majority had returned, albeit with staggered schedules. Undoubtedly, some of the publisher platform data reported here might be lower than it would have been had there not been the pandemic. The same could also be true for usage of the five publishers within EBSCO. To further evaluate the impact of the pandemic shutdowns on the results of this study, the author of this study looked into usage of the 5 publishers across vendor platforms, EBSCO, EJC, and JSTOR for calendar year 2019 and found no major difference in usage. Calendar year 2019 was used, rather than the fiscal year running July 2019–June 2020, because 2019 represents the most recent period when the pandemic was not occurring.[11]

Moreover, in addition to the COUNTER 5 data, UD Libraries also monitors usage of our EBSCO subject databases and our EDS within GoogleAnalytics, and for the past 4 years we have

[10] The drawback with this trend of users selecting articles that are 1-2 years old because they are immediately available within EBSCO is that they might be missing out on the latest research.
[11] For specifics on calendar year 2019 usage, please see the "Appendix 2" section of this study.

not seen a significant decrease in the number of page views (roughly 1.1 million per year) nor in the number of user sessions seen collectively across our EBSCO databases and EDS. Through Google Analytics custom reports, it was determined that 63% of pageviews (or ~660,000 pageviews per year) occurred within UD's configuration of EDS, while 76% of user sessions (or ~105,000 sessions per year) occurred within UD's configuration of EDS. To put this in context, UD's subscription to Web of Science had only 11,175 searches within the same time period. Likewise, the number of linkouts from EBSCO search results to external sources such as JSTOR or SAGE journals has increased over the past two years, suggesting that users are finding more use with the custom links we have configured within the subject databases and in EDS. All of this usage demonstrates the central importance of EBSCO's databases, Full Text Finder, Publication Finder Interface, and discovery layer to the university's instruction and research.

In addition, the data showing the average number of article downloads per title (Figure 6) reveals how JSTOR and EBSCO titles are receiving a far higher average of downloads than titles within the vendor platforms or within OhioLINK's EJC. This likely suggests that these platforms are being used differently, and it's worth speculating that this might indicate that users within JSTOR and EBSCO are relying on the search features more frequently and that perhaps their relevancy rankings are bringing certain titles to the top of the results leading to heavy usage. A simpler way to think about this data point is the mere fact that it is a good trend if a library has a certain set of titles with a far higher average number of full-text retrievals within one platform than the averages recorded for the same publisher in other platforms. Only Springer's vendor platform, SpringerLINK, has a higher average download rate than the average for Springer titles within EBSCO.

There is no denying the importance of the publisher platforms such as SpringerLink and Wiley Online Library at institutions such as the University of Dayton, as can be seen here through the usage data showing direct retrievals, but it's worth suggesting that this only provides one part—perhaps a rather minor part—of the broader usage at an institution like University of Dayton. If the e-resources community focuses on the referral data to the vendor platforms provided by publishers, then the central importance of the publisher platforms can be overstated.

Note that this cautionary comment does not apply to all publishers and vendor platforms. In the case of a select number of publishers, the vendor-controlled platform is often the sole space where full-text articles can be retrieved from, as is the case with Elsevier's ScienceDirect and IEEE Xplore. In these cases, the referral data might indeed be the full story on usage, and if the library referrals are as low as 10%, then it does call for further discussion and concern with these specific publishers. But as is demonstrated in this paper, platforms like IEEE Xplore and ScienceDirect, while widely used, are still not being used more than EBSCO to retrieve full-text, and they seem to be among the minority of publishers who refuse to allow their full-text content to be housed in EBSCO with one- to two-year embargoes. Moreover, if prior studies analyzing vendor-supplied usage data have only examined referral data from this subset of publisher platforms—where the full-text is exclusively available—then they are likely excluding several prominent publishers that nearly all universities subscribe to, namely Wiley, Springer, and SAGE.

Across all publisher platforms discussed in this paper, it might be true that user preferences are changing when it comes to how they interact with the discovery layer and other customized library databases, or it simply might be the case that the trends discussed here have remained consistent over time. Users might prefer results with full text directly available, where

they don't have to rely on a link to an external source or a Full Text Finder window. Likewise, the relevancy ranking provided by EBSCO might be increasing the frequency at which the end user sees these full-text results, over those containing custom links to external platforms (as has been similarly suggested by Breeding in 2015). Ultimately, publishers like IEEE and Elsevier might be missing out on additional usage, citations, and impact across their e-journal publications because their content is not included within EBSCO's full-text holdings.

Based on these findings, the next step is to examine usage data at other institutions in a similar manner to how it is analyzed and visualized here. If usage data from other universities proves that UD is an anomaly, then further studies should consider why this is the case. UD has a longstanding library instruction program with heavy integration across the university's various curricula. We are also one of the most residential campuses in the country with the majority of our undergraduates live in university housing for the entirety of their degrees, ensuring that students remain within the library's IP ranges. Both of these factors could be steering students—and to a lesser extent faculty—toward the library website, discovery layer, and subject databases in heavier numbers than at other institutions.

On the other hand, if this pattern of usage is not an anomaly and it repeats across other institutions, both similar and different from a medium-size research university like the University of Dayton, then these findings could have huge implications for how decisions concerning library discovery layers should be made. Any major decision to abandon a discovery layer (or even repurpose it) based on insights from the referral data alone could have wide-ranging negative consequences for faculty and students who rely on this search platform (and, in the case of EBSCO EDS, for anyone who also relies on its integration with the corresponding subject databases). Based on recent publications and conferences the author of this study has attended, it

often feels like library researchers and leaders are encouraging this change, citing libraries like the one at Utrecht University as an example (Dempsey, 2020, p. 12) or by boldly asserting how users are no longer using or starting with the library anymore (Evans & Schonfeld, 2020, p. 15; Lean Library, 2021, p. 28). As is the case at University of Dayton, the collective EBSCO holdings and search interfaces might be serving as the primary "ecosystem" for users to search for, locate, and retrieve relevant full-text articles *without* having to go to the external publisher platforms.

———

Acknowledgements

The author would like to thank Christina Beis at the University of Dayton Libraries for providing useful feedback in the drafting and editing of this publication.

Data availability statement

All data discussed/analyzed in this study is available in figshare at https://doi.org/10.6084/m9.figshare.17263823.v1.

Declaration of interest statement

The author reports there are no competing interests to declare.

References

Bates, J., Best, P., McQuilkin, J., & Taylor, B. (2017). Will Web Search Engines Replace Bibliographic Databases in the Systematic Identification of Research? *The Journal of Academic Librarianship*, *43*(1), 8–17. https://doi.org/10.1016/j.acalib.2016.11.003

Breeding, M. (2015). Discovery product functionality. Library technology reports, 50(1), 5-32. https://doi.org/10.5860/ltr.50n1

Cummings, J. (2021). Online navigation to journal articles: How are journal articles retrieved by researchers and students at an academic institution; a quantitative examination of HTTP referer [sic] data. *Journal of Electronic Resources Librarianship*, *33*(2), 63–74. https://doi.org/10.1080/1941126X.2021.1911446

Calvert, K. (2015). Maximizing Academic Library Collections: Measuring Changes in Use Patterns Owing to EBSCO Discovery Service. *College & Research Libraries*, *76*(1), 81–99. https://doi.org/10.5860/crl.76.1.81

Dempsey, Lorcan 2020. Foreword to Resource Discovery for the Twenty-First Century Library: Case Studies and Perspectives on the Role of IT in User Engagement and Empowerment, xxi–xxxii. Edited by Simon McLeish. London: Facet Publishing. OCLC Research Foreword Preprint. https://doi.org/ 10.25333/p8j4-1411

Evans, G. & Schonfeld, R.C. (2020). It's Not What Libraries Hold; It's Who Libraries Serve: Seeking a User- Centered Future for Academic Libraries. Ithaka S+R. 23 January 2020.

Evelhoch, Z. (2016). Web-Scale Discovery: Impact on Library Database Web Page Views and Usage. *Journal of Web Librarianship*, *10*(3), 197–209. https://doi.org/10.1080/19322909.2016.1191048

Hayman, R. (2017). Early Career Researchers Demand Full-text and Rely on Google to Find Scholarly Sources. *Evidence Based Library and Information Practice*, *12*(4), 256. https://doi.org/10.18438/B8WD4G

Inger, S., & Gardner, T. (2016). How readers discover content in scholarly publications. *Information Services & Use*, *36*(1–2), 81–97. https://doi.org/10.3233/ISU-160800

Lean Library. (2021). *Librarian Futures*. SAGE Publishing. https://doi.org/10.4135/wp.20211103

Levine-Clark, M., Mcdonald, J., & Price, J. (2014). The effect of discovery systems on online journal usage: a longitudinal study. *Insights: The UKSG Journal*, 27(3), 249–256. https://doi.org/10.1629/2048-7754.153

Ngo, L., Hennesy, C., & Knabe, I. (2019). The Impact of Web-Scale Discovery on the Use of Electronic Resources. *Serials Review*, *45*(4), 227–238. https://doi.org/10.1080/00987913.2019.1695343

Nicholas, D., Boukacem-Zeghmouri, C., Rodríguez-Bravo, B., Xu, J., Watkinson, A., Abrizah, A., Herman, E., & Świgoń, M. (2017). Where and how early career researchers find scholarly information: How ECRs find scholarly information. *Learned Publishing*, *30*(1), 19–29. https://doi.org/10.1002/leap.1087

Nicholas, D., Rowlands, I., Huntington, P., Jamali, H. R., & Hernandez Salazar, P. (2010). Diversity in the e-journal use and information-seeking behaviour of UK researchers. Journal of Documentation, 66 (3), 409–433. https://doi.org/10.1108/00220411011038476O

Robinson, L., Snyder, R., & Burke, D. (2013). Discovery and Analysis of the World's Research Collections: JSTOR and Summon Under the Hood. *The Serials Librarian*, *64*(1–4), 196–199. https://doi.org/10.1080/0361526X.2013.760317

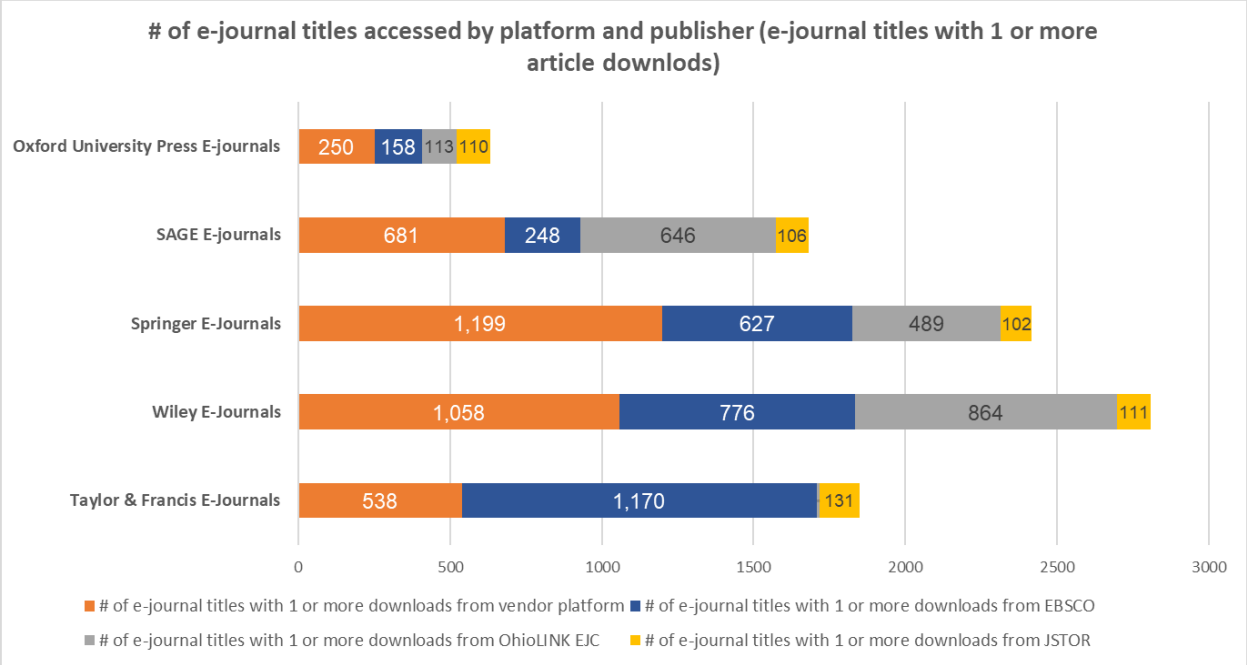
Wang, Y. (2020). Web-scale discovery and Google Scholar: A study in use patterns. *Journal of Electronic Resources Librarianship*, *32*(1), 1–10. https://doi.org/10.1080/1941126X.2019.1709722

Wang, Y., & Howard, P. (2012). Google Scholar Usage: An Academic Library's Experience. *Journal of Web Librarianship*, *6*(2), 94–108. https://doi.org/10.1080/19322909.2012.672067

Appendix 1: Number of E-Journal Titles Used by Publisher & Platform

Publisher	# of e-journal titles with 1 or more downloads from vendor platform	# of e-journal titles with 1 or more downloads from EBSCO	# of e-journal titles with 1 or more downloads from OhioLINK EJC	# of e-journal titles with 1 or more downloads from JSTOR	Total across all (by publisher)
Taylor & Francis E-Journals	538 (out of 755 subscription titles available)	1,170	10	131	1,849
Wiley E-Journals	1,058 (out of 1,899	776	864	111	2,809

	subscription titles available)				
Springer E-Journals	1,199 (out of 2,413 subscription titles available)	627	489	102	2,417
SAGE E-journals	681 (out of 1,164 subscription titles available)	248	646	106	1,681
Oxford University Press E-journals	250 (out of 267 subscription titles available)	158	113	110	631
Total across all (by platform)	3,726	2,979	2,122	560	


# of e-journal titles accessed by platform and publisher (e-journal titles with 1 or more article downlods)
Oxford University Press E-journals
250
158
113
110
SAGE E-journals
681
248
646
106
Springer E-Journals
1,199
627
489
102
Wiley E-Journals
1,058
776
864
111
Taylor & Francis E-Journals
538
1,170
131
0
500
1000
1500
2000
2500
3000
# of e-journal titles with 1 or more downloads from vendor platform
# of e-journal titles with 1 or more downloads from EBSCO
# of e-journal titles with 1 or more downloads from OhioLINK EJC
# of e-journal titles with 1 or more downloads from JSTOR

Appendix 2: Calendar Year 2019 Usage Compared to Data Analyzed in this Study

Vendor Platforms	Total item requests CY19
Wiley Online Library	17,315
Oxford Academic	6,204
SAGE Journals	11,906
Springerlink	13,371
T&F online	5,034
Total across all	**53,830**

EBSCO	Total item requests CY19
Wiley in EBSCO	21,177
Oxford	5,410
Sage	5,776
Springer in EBSCO	8,923

T & F in EBSCO	32,642
Total across all	**73,928**

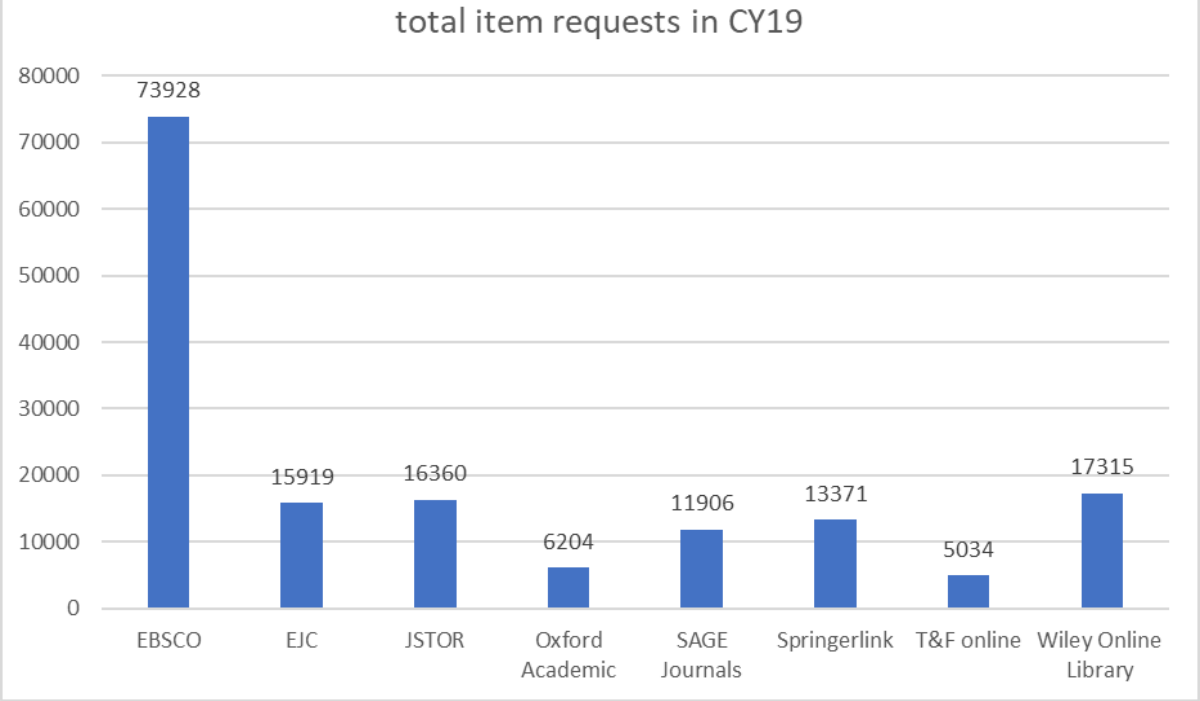
EJC	Total item requests CY19
Wiley in EJC	4,933
Oxford in EJC	539
Sage in EJC	2,390
Springer in EJC	8,008
T&F in EJC	49
Total across all	**15919**

Publishers accessed via JSTOR	total item requests CY19
Wiley	3914
Oxford	4711

Sage	2943
Springer	2257
T&F	2535
Total across all	**16360**

Thinking solely in terms of platforms used, with EBSCO being one of the platforms, here's how this data ends up:

Platform	total item requests in CY19
EBSCO	73928
Wiley Online Library	17315
Oxford Academic	6204
SAGE Journals	11906
Springerlink	13371
T&F online	5034
EJC	15919
JSTOR	16360



Above visual shows the same five publishers by platform used. Note that it's highly similar to how things looked in FY21 (as visualized on the next page), showing consistent usage, despite the pandemic.

Breakdown of full-text retrievals by different platforms