Library Catalog Log Analysis in E-book Patron-Driven Acquisitions (PDA): A Case Study

Cristóbal Urbano, Yin Zhang, Kay Downey, and Thomas Klingler

Patron-Driven Acquisitions (PDA) is a new model used for e-book acquisition by academic libraries. A key component of this model is to make records of e-books available in a library catalog and let actual patron usage decide whether or not an item is purchased. However, there has been a lack of research examining the role of the library catalog as a tool for e-book discovery and use in PDA. This paper presents a case study of using PDA for e-book acquisition in an academic library, with a focus on the role of the library catalog in this purchasing model. The implications and challenges are also discussed.

atron-Driven Acquisitions (PDA), also called Demand-Driven Acquisition (DDA), refers to the automated purchasing of e-books based on patron use. It is a new acquisition model widely adopted for purchasing e-books in academic libraries and has been recognized by the Association of College and Research Libraries (ACRL) as one of the "2012 Top Ten Trends in Academic Libraries."[1] In January 2012, Kent State University Library (KSUL) implemented a pilot initiative for a PDA e-book purchasing model, using a combination of the acquisition services provided by their primary book jobber, Yankee Book Peddler (YBP), and the access services supplied by an e-book distributor (ebrary).

Through this model, KSUL provided access to a predetermined set of e-books to authorized users by making e-book discovery records available in the library catalog, KentLINK. All of these e-books met the library's print book approval guidelines and were deduplicated against library holdings before being uploaded to the catalog. Patron usage of each e-book was tracked by several measures; and, once the cumulative usage of an e-book exceeded a set threshold level, the ebrary system triggered an automatic purchase. A purchase was triggered by one of the following usage measures: 10 page views, 10 view minutes, 1 copy (such as copying text from the e-book), 1 print, or 1 chapter/range download. Short-term loan options were not part of the purchase model

Cristóbal Urbano is Associate Professor at Universitat de Barcelona, Departament de Biblioteconomia i Documentació; e-mail: urbano@ub.edu; Yin Zhang is Professor at Kent State University's School of Library and Information Science; e-mail: yzhang4@kent.edu; Kay Downey is Associate Professor and Collection Management Librarian; and Thomas Klingler is Professor of Bibliography and Assistant Dean for Technical Services in the University Libraries at Kent State University; e-mail: mdowney1@kent.edu, tk@kent.edu.

for the KSUL PDA program. Patrons discovering e-books via the library catalog were unaware of whether a particular PDA e-book was still available for purchase or had already been acquired. In this way, PDA becomes a seamless selection process that augments traditional library approval purchases.

Problem Statement and Objectives

The discovery pool of pre-approved PDA e-book records within the library catalog is a key element in this acquisition model for e-books to be discovered, used, and then triggered for purchase. So far, very few empirical studies have examined how library users interact with the catalog before they connect with the e-books at the provider's platform.

This study is the first to closely explore the role of the library catalog in e-book discovery and use through a comprehensive catalog transaction log analysis of the KSUL PDA pilot project that ran from January 1 to June 30, 2012.[2] The specific questions to be addressed include the following:

1. How do patrons use the catalog to discover e-books?
2. How is catalog use associated with e-book triggering and use?

The answers to these questions will help establish to what extent the catalog is used for PDA e-book discovery and whether or not more frequent e-book use is linked to catalog searches and full bibliographic record displays. These answers are critical to library practices and PDA implementations, as they will help improve the efficacy of the PDA model and influence PDA acquisition policies. By examining actual catalog use and contrasting it with general e-book usage, this study will shed light on possible improvements to library catalogs as e-book discovery tools and provide additional parameters to include in the e-book provider formula for fine-tuning the purchase triggering process. Finally, the results of this study will yield a better understanding of academic library users' information search and use behavior in today's information environment in which users are accustomed to a variety of discovery tools besides library catalogs, such as search engines and direct information portals.

Literature Review

Due to the recent acceleration of e-book acquisitions, the literature devoted to PDA has experienced a growth that can be traced through the papers published on the topic during the last three years, as well as through some papers and books that offer a good general review of the topic.[3]

PDA is a trending topic in LIS literature, but few studies have focused on user interaction with PDA records within library catalogs (or other tools like federated searching or web-scale discovery services), despite the fact that it is an acquisition system that relies on the local catalog as the key discovery tool to access commercial platforms offering PDA titles. Also, although there has been a long history of studies on the use and users of OPACs (Online Public Access Catalogs) since their introduction in libraries, with most research using transaction log analysis methods, this approach has not been used to study the use of the library catalog as a way to access e-books, nor to study PDA programs.[4]

Over the last few years, several studies on e-book use and users reveal a mixed image involving both success and failure of e-book usage in libraries.[5] Some studies have an interesting focus on the ways users discover and access e-books, but few trace the users' path from the OPAC to the commercial provider platform. An interesting and isolated example is the Colorado State University pop-up web survey applied through embedded links from the bibliographic records of a PDA set of e-books.[6] Also, some libraries with uncataloged e-books in their collections have used the statistical reports

from their commercial e-book providers to discover that cataloged titles were actually used more frequently than those that were not cataloged.[7]

The statistical data about e-book use that libraries receive from commercial providers usually follow (more or less) the COUNTER guidelines agreed upon as "de facto" standards for e-resource statistics by libraries and the publishing industry.[8] Generally, however, those reports do not offer information about the "referrer" log field, which could help determine whether the catalog or other library websites are the source of some visits to licensed content. Also, the absence of a user IP address field in these reports makes it impossible to map the uses from different places on campus, or off-campus if user access is via VPN or remote proxy connections. Few studies have undertaken this exploitation of raw log files, because the involvement of a publisher or a consortium is necessary. The few examples of such studies based on raw log files include the internal OCLC Netlibrary study of e-book content and the user study of the OhioLINK Electronic Journal Center.[9]

Previous research based on surveys and other direct methods has suggested that the presence of e-books in the library catalog increases their use.[10] For instance, York University's faculty and graduate students were surveyed to determine how to best promote usage of the library's science e-books, and the authors found that cataloged e-books were accessed more than those that had been promoted by e-mail and other avenues.[11] In similar undertakings at the University of California and the University of Denver, users indicated by 53 percent and 39.3 percent respectively, that the library catalog was their main avenue for the discovery and access of e-books.[12] While these studies paint a picture in which the catalog serves as the main channel through which to access or promote e-books, there are also surveys and other ethnographic studies showing that the catalog, or the library itself in general, is not the main option users choose when beginning their search.[13]

Branding the catalog as a reliable e-book discovery tool has proven challenging. Some libraries with huge e-book collections do not catalog their e-books title by title. This may be because large subscription deals or PDA programs involve significant amounts of investment in cataloging and in catalog maintenance for nonpermanent holdings, creating a cost issue that challenges the catalog as the central tool for e-book access. Consequently, there is ample room for deeply studying the role of metadata in the discovery, selection, and acquisition of PDA e-books, as well as for the analysis of the interaction between catalogs, tools for federated search or web-scale discovery, vendor platforms, and library website guides as methods for e-book discovery and access.[14]

Other studies with a broader scope involving various institutions give us a clearer idea of the challenges libraries face in promoting e-book collections. An ebrary survey of students from different countries found that 3,132 respondents out of 5,886 said that they never use library e-books; and, when asked why, the two highest responses were "I do not know where to find e-books" (1,790) and "I prefer printed books" (1,420).[15] These figures could certainly have changed in the last five years, but librarians who have participated in the 2012 annual *Library Journal* survey on e-book usage have similar concerns about these issues. For example, to the question "What hinders students/faculty from using your library's e-book content?" 52 percent of librarians answered "Unaware of e-book availability" and 32 percent answered "Difficult to find/ discover."[16] In addition, the 2009 HighWire survey on e-books showed that, despite librarians' efforts toward making the catalog a tool to locate e-books, they are concerned that users may consider the catalog very similar to other tools like Google, Amazon, publisher websites, or library websites.[17] This perception challenges the notion that library catalogs serve as the primary discovery tool for PDA e-books.

With the PDA e-book model, e-book management in academic libraries is facing new challenges to library workflow, procedures, and policies, due to different approaches to cataloging practices among a variety of libraries when uploading a vast number of temporary records with uneven quality. Technical services professionals warn that, despite the many papers published on PDA, little research has been done on the new technical services processes that are required by the model's workflow.[18]

Promoting e-book use and undertaking PDA programs has also put online library catalogs in the spotlight again from the point of view of publishers and vendors who need their e-books better exposed to readers.[19] However, this shift happens at the same time that the debate about the future of the catalog is gaining momentum in the discussion of broader topics like Web searching, metadata harvesting, and discovery tools.[20] What is interesting to observe is that PDA for e-books relies on the library catalog as a key component for discovery; yet, due to competition from web search engines and other social web tools, library catalogs are at a crossroads, with change appearing inevitable. Some believe that the role of the catalog as a discovery tool in this web-centric scenario is in crisis because libraries should focus on the delivery side of the work rather than on the discovery role, while others express support for the next generation of library discovery tools or for future catalogs to be shaped as new social web spaces.[21]

In summary, it is important to note that PDA analysis of catalog performance needs to be viewed in the context of Google-influenced user behavior in which users tend to prefer Internet search engines over the library catalog due to the perceptions of inefficient search, browsing, and display functions in the latter.[22] According to recent Ithaka S+R reports on information seeking of academic users in the United States, electronic research resources are preferred, and discovery tools such as Google Scholar and Microsoft Academic Search play a key role in supporting the discovery of scholarly materials. Of importance is that, according to the same Ithaka reports, users have a more library-centered approach when they search for a known item they have discovered elsewhere, and U.S. academic library directors desire a strategy for e-resource discovery that will keep the library relevant.[23]

In general, this user behavior has not only had an important impact on the decline of the catalog as a means of discovery; it has also resulted in the increasing predominance of simple keyword queries in the library catalog itself. According to a study by Lau and Goh, keyword searches contributed to 68.9 percent of all queries, while other options such as title, author, and subject only accounted for 16.5 percent, 8.2 percent, and 6.4 percent of all queries, respectively.[24] To accommodate user search preferences, libraries have attempted to set their catalog's default search to use basic keyword functionality, and most libraries have placed their web-scale discovery service search form as the main option at their home page for users to explore their various collections. This trend should also be considered in the future for PDA assessment.[25]

Methods

In this case study, patrons could discover PDA e-books only via either the library catalog (KentLINK) or through the direct links to the ebrary site (for instance from LibGuides, A–Z listings, or e-reserves), as during the e-book pilot project KSUL offered no access to their collections through other means such as federated search tools or Web-scale discovery services. The methodological approach chosen was catalog log analysis. This method was used because the commercial e-resource providers—ebrary in this case—do not supply information about the referral sources that bring users to their hosting servers in their standard statistical reports. Therefore, the only unobtrusive research method to examine how KSU users reached PDA e-books was through the analysis of KSUL catalog server transactions logs.

Data Sample and Data Sources

The sample of e-book discovery records that were uploaded by January 3, 2012, and remained throughout the study included 20,062 e-books. This sample contained e-books published in 2009 or later that met the KSUL YBP approval plan profile. Among those in the e-book sample, 518 (2.6%) were purchased through PDA during the pilot period. All the triggered titles were purchased under single user licenses that do not grant access for simultaneous users.

The data sources included: 1) the KentLINK and YBP ordering data for the 20,062 e-book records in the discovery pool; 2) the ebrary Trigger Report during the PDA pilot that contained details regarding which e-books were triggered and usage measures up to the triggering point; 3) the catalog log data extracted from the Apache Web server logs recorded for the KentLINK website for the entire year of 2012; and 4) the ebrary Title Report that served as the source for tracking the actual use of e-books for the entire year of 2012. The full year of 2012 was chosen to track e-book usage during the two main academic terms at the university, leaving almost six months of usage time for titles triggered toward the end of the pilot period in June 2012. What is important about the Title Report is that it offers information on the use of the nonpurchased PDA e-books as well as those that were triggered.

Activity Recorded in Catalog Server Transaction Logs

KentLINK server transaction logs track the clicks on the link in the full bibliographic record toward an e-book at the ebrary platform (see table 1). User searches that occur without clicks on these links are not tracked. For the purpose of this study, various e-book access approaches are differentiated and defined as below:

- A "full-orthodox" access is defined as an entry to the ebrary portal (and subsequent trigger action) that results from a user accessing an e-book via links in the full bibliographic catalog record alone. The full record display offers more details about the e-book. Only the full-orthodox access to ebrary from the catalog can be completely traced in the library transaction logs. We hypothesize that requesting an e-book from its full bibliographic record display is an indication of a more conscious, individual e-book request from the user.
- A "not full-orthodox" access to the e-book is defined as one of the following two situations: 1) when the user clicks on the link to the e-book from the catalog intermediate search results brief display; or 2) when the user interacts directly with an ebrary e-book without a direct link from any catalog display (full or brief). Users can reach the e-book without a KentLINK catalog referral after an access coming from the catalog because, once inside the ebrary portal, the user can continue browsing or perform another search. Users can also go directly to the portal from their bookmarks, links suggested by colleagues or teachers, or from a library guide like the "E-books Online Reference Shelf" offered through the KSUL website.

Analysis of Catalog Logs and ebrary Stats Reports

We obtained the catalog record ID for the full list of 20,062 PDA e-books to trace them to the KentLINK log file. For the 518 e-books purchased through PDA during the pilot period, we checked logs on the exact date of triggering to know whether the triggering process for purchasing was linked to an actual search in the catalog.

A batch program was written to extract all of the catalog log lines where a record ID of one of the ebrary e-books appears in the "Request" field. Every time a request for an ebrary book is made from the full bibliographic record in KentLINK, the log file records two lines, which include the bibliographic record ID and the search type

TABLE 1				
Sample Catalog Transaction Logs Tracing a Click of an E-book Access Link within the Full Bibliographic Record				
Port	IP	Month/ Day/Time	Request (for an e-book with a bibliographic record ID as **b4087065**)	Referrer
80	70.xxx. xx.29	01/11/ 13:39:06	GET /search~S1?/ tOedipus+Rex/ toedipus+rex/101,108,108,E/ l9xx~**b4087065**&FF=toedipu s+rex+english&108,108,,1,0/ indexsort=c HTTP/1.1	http://kentlink.kent.edu/ search~S1?/tOedipus+Rex/to edipus+rex/101%2C108%2C 108%2CE/frameset&FF=toe dipus+rex+english&108%2C 108%2C/indexsort=c
80	70.xxx. xx.29	01/11/ 13:39:06	GET /search~S1?/ tOedipus+Rex/toedipus+rex/ 101%2C108%2C108%2CE/li nk~**b4087065**&FF=toedipus +rex+english&108%2C108% 2C/indexsort=c HTTP/1.1	http://kentlink.kent.edu/ search~S1?/tOedipus+Rex/ toedipus+rex/101,108,108,E/ l9xx~b4087065&FF=toedipu s+rex+english&108,108,,1,0/ indexsort=c

and query (see table 1). We considered a period of 30 minutes to be an indication of timeout within the search session in the catalog. To define the unique clicks to an e-book, we considered this timeout as well as any changes for each record ID in the log line representing a different value for IP, request, referrer, agent, or port.

To establish when a search in the catalog matches with a trigger action on a given date, we filtered the logs by IP address. Only IPs from the KSU range, which included VPN connections outside the campus, could trigger the purchase, due to the IP recognition authentication that occurs when the user follows the link from the catalog to the ebrary content. Users with external IPs could not access the full-text content in the ebrary KSU dedicated link or trigger the purchase by following the link from the catalog.

We matched the ebrary Trigger Report with the logs to trace whether a full-orthodox access had been performed from a KSU IP. Since the ebrary report offers only the date and not the time of the triggering, when more than one full-orthodox access occurred on the trigger date we considered all the searches performed that day as potential clicks for triggering and analyzed them in a separate set from the cases with a unique IP/search as a triggering agent.

Results
1. How do patrons use the catalog to discover e-books?
For the PDA sample pool of 20,062 discovery records, there were 3,254 "full-orthodox" clicks from the catalog to the ebrary content. These 3,254 clicks represent a first set of 2,369 different unique searches from any IP: that is, searches performed for each different record ID, with a different IP, within a timeout window of 30 minutes and with a different search query. If we look only at searches performed from KSU IPs over the sample pool, there were 1,822 searches and a subset of 1,098 searches over the triggered titles.

The distribution of the searches for the sample pool of PDA titles and for the subset of e-books purchased after user activity triggered the acquisition is detailed in table 2. The first column presents the type of search and the other features that users could use as part of their search activities, which include filtering the results to retrieve only e-resources/e-books, using a mobile device for the search, or searching from a personal

TABLE 2
Types of Searches Performed Before Requesting an E-book Access Link through the Full Bibliographic Record in the Catalog

Type of Search	Sample Pool (any IP)	%	Sample Pool (KSU IP)	%	Triggered E-books (KSU IP)	%
Author	59	2.49%	42	2.31%	24	2.19%
Record ID	88	3.71%	74	4.06%	54	4.92%
Boolean	33	1.39%	26	1.43%	8	0.73%
Subject	111	4.69%	79	4.34%	20	1.82% *
ISBN	56	2.36%	46	2.52%	43	3.92%
Other Standard Numbers	4	0.17%	3	0.16%	1	0.09%
Title	345	14.56%	269	14.76%	172	15.66%
General Keyword	1,664	70.24%	1,277	70.09%	774	70.49%
Keyword in Subject	1	0.04%	0	0.00%	0	0.00%
Keyword in Title	8	0.34%	6	0.33%	2	0.18%
TOTAL	2,369	100.00%	1,822	100.00%	1,098	100.00%
Other Factors and Features						
Filter by E-resource/ E-book	184	7.77%	105	5.76%	51	4.64% *
Mobile Devices (Tablets & Phones)	31	1.31%	17	0.93%	11	1.00%
Authenticated Users [port 443 (https)]	120	5.07%	84	4.61%	45	4.10%
Searches That Return >50 Records	976	41.20%	763	41.88%	420	38.25%
* = significant difference of proportions at 95% confidence level.						

catalog user account under a secured web server connection. The data reported for each type of search is presented in the next three columns representing successive subsets of all searches: searches performed from any IP, searches performed only from a KSU IP, and finally any search linked to the triggered e-books. Users with non-KSU IP addresses included in the first column can search the catalog but can't access the ebrary KSU dedicated entrance or trigger any e-book purchase.

In general, these data are consistent with the low elaboration of searches performed by patrons in library catalogs, similar to what is done in Web search engines. That is, among the numerous retrieved records, only the first ones are picked from the initial results page. Predictably, keyword search was the most prevalent search type used, followed by title search. The keyword search form is the default option for the OPAC and for the search window embedded on the KSUL home page. The distribution of different search types is similar to what was found in other studies involving general OPAC use.[26]

The comparison of the types of searches for the subset of triggered e-books against the sample pool of PDA e-books indicated consistency with only some minor differ-

TABLE 3
**Selection of Full Bibliographic Records from Keyword and Boolean Searches
That Returned More Than 50 Records**

Selection of Records in Keyword and Boolean Search Sets > 50	Sample Pool (any IP)	%	Sample Pool (KSU IP)	%	Triggered E-books (KSU IP)	%
Total Searches That Return >50 Records	976	100.00%	763	100.00%	420	100.00%
Selection First Record of the Set	327	33.50%	255	33.42%	188	44.76% *
Selection Records 2 to 5 of the Set	270	27.66%	211	27.65%	105	25.00%
Selection Records 6 to 50 of the Set	361	36.99%	282	36.96%	126	30.00% *
Selection Records Beyond 50TH	18	1.84%	15	1.97%	1	0.24% *
* = significant difference of proportions at 95% confidence level.						

ences. We can see that, for the triggered e-books, some types of searches are slightly lower (Boolean, subject, and filter by e-resource) and others are slightly higher (record ID and ISBN). These differences are only significant in two cases, and assessment of the statistical significance of the difference of shares in those three sets was done through the calculation of confidence intervals for proportions at 95 percent confidence level. The proportions of cases in the "triggered e-book" set are significantly lower than the proportions in one of the other sets only in the cases of "Subject" and "Filter by e-resource/e-book."

Since a great proportion of keyword and Boolean searches retrieve a large number of records, it is interesting to see the degree to which users browsed and scrolled the list of results for these sets before they clicked the e-book link. Also, the fact that a great number of keyword searches include all the words of a title could explain the high numbers of first record choice. KentLINK's default intermediate record display includes up to 50 brief records, and, at first glance and with a little scrolling, one can reach up to five records. We have analyzed the search results under these two thresholds in table 3. The noted difference among the three sets is that the selection of the first record is significantly more prevalent in triggered e-books. The main explanation of such a difference is that "title-like" searches for known items through the keyword search could be more common for the triggered e-books, but more research is needed to explore this further. It appears that users tend to search the catalog for known items, looking to check a title's presence in the collection whether by a title query or by a keyword search using the default simple search box.[27]

It was observed from this study that the number of searches filtered by document type "e-resource/e-book" is very low (4.64 percent for the triggered set), which could be due to problems with the search form layout or catalog features, or the lack of user awareness about the availability of e-books in the catalog.[28] This low level of filtering searches by "e-resource/e-books" may suggest that users do not generally use the catalog to search for e-books exclusively.

TABLE 4					
E-Books Used at Least Once without Traces in the Catalog Logs					
	Total	**Used**	**% Used**	**Used but not Traced in Catalog**	**%**
Sample Pool	20,062	1,131	5.64%	139	12.29%
Triggered E-books	518	518	100.00%	114	22.01%

2. How is catalog use associated with e-book triggering and use?

The ebrary Title Report for 2012 informs us not only about the volume of use, but also about which e-books actually experienced at least one visit. Only a small fraction of the 20,062 e-books from the sample were used, and an even smaller fraction triggered, during the pilot test (see table 4). We can see that the share of triggered e-books that have been used without any link to a catalog search on any date is proportionally higher than for the complete PDA set: 22.01 percent of the triggered books have never been accessed from a full catalog record, against 12.29 percent for the general sample pool.

To assess whether the type of catalog activity has some kind of relationship with the average level of use for the 518 triggered e-books, the use activities of these e-books have been analyzed and summarized in table 5. Table 5 has two sections: one for the 373 titles triggered via "full-orthodox" triggering on the trigger date, and the other for the 145 titles that do not match any "full-orthodox" catalog search on the trigger date. This second section has also been split further between the 114 titles with no trace at all in the logs from KSU IPs, and the 31 titles with a trace but not on the triggering date.

It can be observed that the average use for the "full-orthodox" triggered e-books is greater than for the "not full-orthodox." These 31 titles have the highest level of users per session, as well as the highest number of page views. For the three indicators of use as measured by average user session, average page view, and average page views per user, the books with some presence in the catalog logs have a greater frequency of use with higher average page views per user session. However, the differences are small, with a range from 14.00 to 14.56 pages per user session—low figures that suggest shallow reading behavior.

As shown in table 6, the number of access attempts of e-books from the catalog before triggering has a strong relationship with the frequency of e-book use during the entire year of 2012. Those e-books that received two or more visits from full bibliographic

TABLE 5					
E-Book Triggering Actions and Associated E-Book Use					
	Titles	**% of Titles**	**Average User Sessions**	**Average Page Views**	**Average Page Views per User**
Total Triggered E-books	518	100.00%	5.46	90.00	14.43
"Full-orthodox" Triggering	373	72.01%	5.76	93.97	14.56*
"Not full-orthodox" Triggering	145	27.99%	4.69	79.77	14.08
> No Catalog Logs from a KSU IP	114	22.01%	3.80	53.52	14.00
> In Catalog Logs from a KSU IP but not Matching with a Trigger Date	31	5.98%	7.97*	176.32*	14.38
* = Highest value under each column					

TABLE 6
Click-throughs to E-Books in Catalog Prior to Triggering and Associated E-Book Use

Clicks-through Prior to Triggering	Titles	Average User Sessions	Average Page Views	Average Pages per User
0	132	4.34	66.30	14.06
1	283	4.25	65.30	13.44
2	62	8.03	156.00	17.09
3	27	7.44	138.59	19.99*
4 to 10	14	25.21*	426.57*	15.37
Total Triggered	518	5.46	90.00	14.43
* = Highest value under each column				

record links (up to the trigger date) had higher levels of use during 2012. This result raises the question of whether or not avoiding "first shot" triggering upon initial use could be an improvement for some PDA models.

Additionally, we also assessed e-book use by type of search for those titles with a triggering point clearly linked to a catalog search. Table 7 summarizes the average usage for the 361 titles that were retrieved through one specific search type at the triggering point.

The results show that e-books triggered through author, subject, title, or keyword searches have similar levels of use. It should be noted that many keyword searches are "title-like" searches (with complete book titles using the default search box for the queries). Such keyword searches linked to actual triggering actions have a similar level of use as title searches. A similar pattern was also observed for "author-like" searches. We suggest further analysis of keyword searches to focus on those that are more subject-oriented to understand the effect of different search types.

Without considering ISBN searches (because the few cases are not relevant enough), the highest average number of user sessions is linked to author search, and the highest number of page views corresponds to title search. These findings indicate less vague needs from the users, who likely have some knowledge about the books they are looking for, which makes user-driven acquisition less random.

TABLE 7
E-Book Use by Catalog Search Type for E-Books Retrieved through One Search Type on the Trigger Date

Type of Searches	Titles	Average User Sessions	Average Page Views	Averages Pages per User
Author Search	10	5.50	78.00	14.40
Bib ID	14	4.21	48.29	11.55
Boolean*	4	4.00	47.25	11.13
Subject	9	5.22	93.56	14.36
ISBN*	2	59.00	631.50	6.64
Title	60	5.32	95.52	13.88
Keyword	262	5.41	88.58	14.84
* = not enough data to be meaningful.				

Discussion

This study examined the role of library catalogs in e-book discovery and use. Specifically, it examined the extent of catalog use as the gateway to access PDA e-books at the vendor site, and how catalog use is associated with e-book triggering and e-book use. The results of this study confirm that full catalog interaction matters in PDA, with more than 70 percent of the purchased e-books being purchased after a trigger action linked with a catalog search and a click-through to the e-book from a full bibliographic record. In addition, the use of e-books was always slightly higher when catalog use was involved. The results also raise some new issues and challenges to be discussed below.

Challenges to the PDA Model

Among PDA e-book purchases, 28 percent occurred without full-orthodox access via a catalog record (see table 5). This means that users triggered the purchases either without seeing the full bibliographic details or directly from the ebrary portal. Users who triggered e-book purchases in this fashion lacked the context of the full KSUL collection. Although we do not have any benchmark studies to assess if this 28 percent bypass level should be considered high, low, or reasonable, KSUL librarians see it as large enough to be considered a real challenge to collection development principles. If users do not avail themselves of the full collection investigation of a catalog search, then PDA could become a limitation rather than added value. Since all discovery records meet collection guidelines, all triggered purchases are relevant to the collection; but, without collection context, bypass purchases could compromise the library's service to its patrons by resulting in spotty collection purchases. One way to control this handicap would be to allow trigger purchases from the catalog only. Further examination of these data could influence best practice and future PDA standards for discovery.

To increase e-book use and awareness, some researchers recommend that libraries should improve their websites with a more holistic approach by offering different pathways and A–Z listings to e-book collections alongside the catalog.[29] Such an approach should be taken with caution when a library has a PDA program with a commercial provider, because it could lead to purchases triggered without the library collection context offered by the library catalog. Additionally, when taking advantage of Web-scale discovery services like Summon, Primo, or EBSCO Discovery Service for e-book discovery and use, it is very important that these platforms have a comprehensive collection of metadata from the local catalog (also PDA records from different providers) to give users a full picture of the collection.[30]

PDA purchasing of two or more e-books by the same user in a short period of time and without using the full bibliographic record in the catalog is a real challenge that needs further study. Can simple known-item searching explain this behavior? Data about e-books used at least once within a year (see table 4) shows that the percentage of used e-books without any trace in the catalog logs is higher (22.01%) among triggered e-books than those in the general e-book discovery pool (12.29%). Do these data simply show that higher-demand titles are subject to both "traceless" use and trigger events? These figures are consistent with the qualitative approach we used for cases of multiple triggering observed in the detailed analysis of logs.

Influence of the Catalog on the Frequency of E-book Use

Findings show that e-book use is higher when a "full-orthodox" triggering from the catalog is involved. However, in our study, the time elapsed between the triggering point and usage tracking is still too short to see if search types and other factors of catalog use make a difference in the level of e-book use. We studied usage data for the entirety of 2012, which means that all of the triggered e-books were available from six

months to less than a year after triggering. Also, because e-book triggering occurred in the spring term at the university, some triggered e-books might have been linked to spring courses not taught during the fall semester. It would be interesting to follow the evolution of the usage for those 518 titles in the years ahead and to compare the usage patterns in recurrent terms.

Those e-books retrieved more than two times via their full bibliographic records before purchase triggering had higher usage afterward (see table 6). These figures could be worth studying more deeply to explore whether the agreement with the vendor should be modified to avoid purchasing e-books in PDA programs upon a "first shot" from only one user that triggers the process after reaching a certain usage threshold. Additionally, it is unclear whether the current PDA implementation with free use until the triggering point is better than other models that do not trigger with first shots, but rather grant access under payment for short-term loans. Data observed in table 6 suggest that e-books with at least four initial uses via full bibliographic records prior to triggering tend to have significantly higher usage overall. Some PDA model scenarios based on longitudinal usage data could help shed light on which PDA implementation works best at the library.

Further research is needed to study the behavior of users after they locate an e-book in the catalog, because, in academic culture, time for reading used to be considered as different from time for searching. In e-book platforms like ebrary that do not have a checkout system of lending to reserve the book, the single-user license titles could force the user that would prefer to have one session of thorough reading to perform at least two (and perhaps many more) sessions. Despite this fact, previous research suggests the level of "immersion in reading" for academic e-books is low—most users use only small portions of e-books during brief sessions.[31] Data from this study show that other usage options like copy/print pages or chapter downloads are very low for all titles. Further analysis is underway to characterize the usability and the features of the platform to understand the data in the search context.

Conclusions

This study's findings show that patrons use the catalog in various ways to discover and use e-books. General keyword and title searches are by far the most popular ways for patrons to find e-books in the catalog. Additionally, e-book usage is higher when an e-book purchase is triggered after consultation with the full bibliographic record. Our findings also show that e-books used via the full bibliographic record more than two times before a triggered purchase have higher usage afterward. These results suggest that the library catalog is a useful e-book discovery tool and that a full bibliographic catalog record consultation is associated with higher usage. However, given the scope of this case study and the fact that there could be other e-book discovery tools besides the library catalog, future studies of other e-book PDA programs with longer time frames are needed to verify that the results are conclusive.

OPAC interface configuration and usability appear to be key issues: whether PDA e-books can be accessed or not from a brief intermediate results display appears to drive the user to a more conscious choice. As title and author searches express very specific and clear needs when the user searches the catalog, improving the search interface and results clustering for a more meaningful and helpful display based on the search query could significantly affect the user's progress toward a PDA trigger event.

While this study illustrates that the full catalog record referrer matters in the PDA program, a considerable share of almost 30 percent of the purchased e-books were triggered without full catalog involvement. An e-book triggered directly in a publisher or aggregator portal without prior interaction with the local catalog may not necessarily

be a justified purchase when there are other similar books in the catalog that could fit the user needs better or just as well. Future research on PDA needs to look into the usability of library catalogs to facilitate e-book discovery and access.

Finally, this study also points to an important area for future research: the users' experiences that do not originate from the catalog and their subsequent effects on e-book discovery and use. Further exploration in this area will help identify possible improvements of current PDA implementations and produce more intuitive and effective approaches for users to discover and use e-books.

Acknowledgments

Notes

1. ACRL Research Planning and Review Committee, "2012 Top Ten Trends in Academic Libraries," *College & Research Libraries News* 73, no. 6 (2012): 311–20; ACRL Research Planning and Review Committee, Environmental Scan 2013 (Chicago: ACRL, 2013), available online at www.ala.org/acrl/sites/ala.org.acrl/files/content/publications/whitepapers/EnvironmentalScan13.pdf [accessed 2 February 2014].

2. The KSU Pilot has been presented in: Kay Downey, "Technical Services Workflow for Book Jobber-Mediated Demand Driven Ebook Acquisitions," *Technical Services Quarterly* 31, no. 1 (2014): 1–12.

3. Barbara Blummer and Jeffrey Kenton, "Best Practices for Integrating E-Books in Academic Libraries: A Literature Review from 2005 to Present," *Collection Management* 37, no. 2 (2012): 65–97; Rajendra Kumbhar, "E-Books: Review of Research and Writing during 2010," *The Electronic Library* 30, no. 6 (2012): 777–95; Merinda McLure and Amy Hoseth, "Patron-Driven E-Book Use and Users' E-Book Perceptions: A Snapshot," *Collection Building* 31, no. 4 (2012): 136–47; Rebecca Schroeder and Tom Wright, "Electronic Books: A Call for Effective Business Models," *New Library World* 112, no. 5/6 (2011): 215–21; *Patron-Driven Acquisitions: History and Best Practices*, ed. David A. Swords (Berlin: De Gruyter, 2011); William H. Walters, "Patron-Driven Acquisition and the Educational Mission of the Academic Library," *Library Resources & Technical Services* 56, no. 3 (2012): 199–213.

4. Christine L. Borgman, "Why Are Online Catalogs Hard to Use? Lessons Learned from Information-Retrieval Studies," *Journal of the American Society for Information Science* 37, no. 6 (1986): 387–400; Christine L. Borgman, "Why Are Online Catalogs Still Hard to Use?" *Journal of the American Society for Information Science* 47, no. 7 (1996): 493–503; Christine L. Borgman, Sandra G. Hirsh, and John Hiller, "Rethinking Online Monitoring Methods for Information Retrieval Systems: From Search Product to Search Process," *Journal of the American Society for Information Science* 47, no. 7 (1996): 568–83; Bernard J. Jansen, "Search Log Analysis: What It Is, What's Been Done, How to Do It," *Library Information Science Research* 28, no. 3 (2006): 407–32; Eng Pwey Lau and Dion Hoe-Lian Goh, "In Search of Query Patterns: A Case Study of a University OPAC," *Information Processing & Management* 42, no. 5 (2006): 1316–29; Aphrodite Malliari, Mersini Moreleli-Cacouris, and Konstantinos Kapsalis, "Usage Patterns in a Greek Academic Library Catalogue: A Follow-Up Study," *Performance Measurement and Metrics* 11, no. 1 (2010): 47–55; Heather L. Moulaison, "OPAC Queries at a Medium-Sized Academic Library: A Transaction Log Analysis," *Library Resources & Technical Services* 52, no. 4 (2008): 230–38.

5. Rusty Kimball, Gary Ives, and Kathy Jackson, "Comparative Usage of Science E-Book and Print Collections at Texas A&M University Libraries," *Collection Management* 35, no. 1 (2009): 15–28; Robert Slater, "Why Aren't E-Books Gaining More Ground in Academic Libraries? E-book Use and Perceptions: A Review of Published Literature and Research,"*Journal of Web Librarianship* 4, no. 4 (2010): 305–31; Jeff Staiger, "How E-Books Are Used,"*Reference & User Services Quarterly* 51, no. 4 (2012): 355–65.

6. McLure and Hoseth, "Patron-Driven E-Book Use," 136–47.

7. Nancy Sprague and Ben Hunter, "Assessing E-Books: Taking a Closer Look at E-Book Statistics," *Library Collections, Acquisitions, and Technical Services* 32, no. 3/4 (2008): 150–57.

8. COUNTER, *Code of Practice for Books and Reference Works Release 1* (2006), available online

at www.projectcounter.org/cop/books/cop_books_ref.pdf [accessed 2 February 2014]; COUNTER, *Code of Practice for e-Resources: Release 4* (2012), available online at www.projectcounter.org/r4/ COPR4.pdf [accessed 2 February 2014].

9. Lynn Silipigni Connaway and Clifton Snyder, "Transaction Log Analyses of Electronic Book (E-Book) Usage," *Against the Grain* 17, no. 1 (2005): 85–89; Paul Huntington et al., "Obtaining Subject Data from Log Files Using Deep Log Analysis: Case Study OhioLINK," *Journal of Information Science* 32, no. 4 (2006): 299–308.

10. Blummer and Kenton, "Best Practices," 65–97.

11. Rajiv Nariani, "E-Books in the Sciences: If We Buy It Will They Use It?" *Issues in Science and Technology Librarian* (Fall 2009), doi:10.5062/F4WS8R5G.

12. Chan Li et al., *UC Libraries Academic E-Book Usage Survey Report: Springer E-Book Pilot Project* (University of California Libraries, 2011), available online at www.cdlib.org/services/uxdesign/ docs/2011/academic_ebook_usage_survey.pdf [accessed 2 February 2014]; Michael Levine-Clark, "Electronic Book Usage: A Survey at the University of Denver," *portal: Libraries and the Academy* 6, no. 3 (2006): 285–99.

13. Ross Housewright, Roger C. Schonfeld, and Kate Wulfson, *Ithaka S+R US Faculty Survey 2012* (New York: Ithaka S+R, 2013), available online at www.sr.ithaka.org/research-publications/ us-faculty-survey-2012 [accessed 2 February 2014]; Cathy De Rosa et al., *Perceptions of Libraries 2010: Context and Community* (Dublin, OH: OCLC, 2010), available online at http://oclc.org/content/dam/ oclc/reports/2010perceptions/2010perceptions_all.pdf [accessed 2 February 2014]; Lynn Silipigni Connaway, Donna Lanclos, and Erin M. Hood, "I Find Google a Lot Easier than Going to the Library Website: Imagine Ways to Innovate and Inspire Students to Use the Academic Library," in *Imagine, Innovate, Inspire: Proceedings of the ACRL 2013 Conference* (Chicago: Association of College and Research Libraries, 2013), 289–300, available online at www.ala.org/acrl/sites/ala.org. acrl/files/content/conferences/confsandpreconfs/2013/papers/Connaway_Google.pdf [accessed 2 February 2014].

14. Ken Chad, *Patron Driven Acquisitions (PDA) and the Role of Metadata in the Discovery, Selection and Acquisition of Ebooks: Final Report for the JISC* (KenChad Consulting, 2011), available online at http://ebmotmet.wikispaces.com/Report [accessed 2 February 2014]; Andrea Dinkelman and Kristine Stacy-Bates, "Accessing E-Books through Academic Library Web Sites," *College & Research Libraries* 68, no. 1 (2007): 45–58; William H .Walters, "E-Books in Academic Libraries: Challenges for Discovery and Access,"*Serials Review* 39, no. 2 (2013): 97–104.

15. *Ebrary 2008 Global Student E-Book Survey* (Palo Alto, Calif.: Ebrary, 2008), available online at www.ebrary.com/corp/collateral/en/Survey/ ebrary_student_survey_2008.pdf [accessed 2 February 2014].

16. LJ/SLJ Library Research Syndicate, *2012 Ebook Usage in U.S Academic Libraries: Third Annual Survey* (Library Journal, 2012), available online at www.thedigitalshift.com/research/ebook-usage-reports/academic/ [accessed 2 February 2014].

17. Michael Newman, *HighWire Press 2009 Librarian eBook Survey* (Palo Alto, Calif.: HighWire Press, 2009), available online at http://highwire.stanford.edu/ PR/HighWireEBookSurvey2010.pdf [accessed 2 February 2014].

18. Jacqueline Belanger, "Cataloguing E-Books in UK Higher Education Libraries: Report of a Survey," *Program: Electronic Library and Information Systems* 41, no. 3 (2007): 203–16; Daniel C. Draper, "Managing Patron-Driven Acquisitions (PDA) Records in a Multiple Model Environment," *Technical Services Quarterly* 30, no. 2 (2013): 153–65; Roman S. Panchyshyn, "Asking the Right Questions: An E-Resource Checklist for Documenting Cataloging Decisions for Batch Cataloging Projects," *Technical Services Quarterly* 30, no. 1 (2013): 15–37; Annie Wu and Anne M. Mitchell, "Mass Management of E-Book Catalog Records," *Library Resources & Technical Services* 54, no. 3 (2010): 164–74; Magdalini Vasileiou, Jennifer Rowley, and Richard Hartley, "Metadata and Providing Access to E-Books," *British Journal of Educational Technology* 44, no. 3 (2013): 518–28.

19. Joseph J. Esposito, "The Problem of Discovery for Patron-Driven Acquisitions (PDA)," *The Scholarly Kitchen* (blog), June 12, 2012, available online at http://scholarlykitchen.sspnet.org/ 2012/06/12/the-problem-of-discovery-for-patron-driven-acquisitions-pda/ [accessed 2 February 2014].

20. Chad, *Patron Driven Acquisitions (PDA) and the Role of Metadata in the Discovery, Selection and Acquisition of Ebooks*; Andrew D. Asher, Lynda M. Duke, and Suzanne Wilson, "Paths of Discovery: Comparing the Search Effectiveness of Ebsco Discovery Service, Summon, Google Scholar, and Conventional Library Resources," *College & Research Libraries* 74, no. 5 (2013): 464–88.

21. Simone Kortekaas, "Thinking the Unthinkable: A Library without a Catalogue: Reconsidering the Future of Discovery Tools for Utrecht University Library," *LIBER Blog*, September 4, 2013, available online at www.libereurope.eu/blog/thinking-the-unthinkable-a-library-without-a-catalogue-reconsidering-the-future-of-discovery-to [accessed 2 February 2014]; Judy Luther and Maureen C. Kelly, "The Next Generation of Discovery," *Library Journal* 136, no. 5 (2011): 66–71;

Lisa O'Hara, "Collection Usage Pre- and Post-Summon Implementation at the University of Manitoba Libraries," *Evidence Based Library and Information Practice* 7, no. 4 (2012): 25–34; Laurel Tarulli and Louise F. Spiteri, "Library Catalogues of the Future: A Social Space and Collaborative Tool?" *Library Trends* 61, no. 1 (2012): 107–31.

22. Jia Mi and Cathy Weng, "Revitalizing the Library OPAC: Interface, Searching, and Display Challenges," *Information Technology and Libraries* 27, no. 1 (2008): 5–22.

23. Housewright, Schonfeld, and Wulfson, *Ithaka S+R US Faculty Survey 2012*; Matthew P. Long and Roger C. Schonfeld, *Ithaka S+R Library Survey 2010: Insights from U.S. Academic Library Directors* (New York: Ithaka R+S, 2010), available online at www.sr.ithaka.org/research-publications/us-library-survey-2010 [accessed 2 February 2014].

24. Lau and Goh, "In Search of Query Patterns," 1316-29; Moulaison, "OPAC Queries," 230–38.

25. Cory Lown, Tito Sierra, and Josh Boyer, "How Users Search the Library from a Single Search Box," *College & Research Libraries* 74, no. 3 (2013): 227–42.

26. Lau and Goh, "In Search of Query Patterns," 1316–29.

27. Kortekaas, "Thinking the Unthinkable;" Malliari, Moreleli-Cacouris, and Kapsalis, "Usage Patterns," 47–55.

28. Belanger, "Cataloguing E-Books," 203–16; Erin Dorris Cassidy, Michelle Martinez, and Lisa Shen, "Not in Love, or Not in the Know? Graduate Student and Faculty Use (and Non-Use) of E-Books," *Journal of Academic Librarianship* 38, no. 6 (2012): 326–32.

29. Dinkelman and Stacy-Bates, "Accessing E-Books," 45–58.

30. O'Hara, "Collection Usage Pre- and Post-Summon Implementation," 25–34.

31. Levine-Clark, "Electronic Book Usage," 285–99; Pervaiz Ahmad and Mark Brogan, "Scholarly Use of E-Books in a Virtual Academic Environment: A Case Study," *Australian Academic and Research Libraries* 43, no. 3 (2012): 189–213; Connaway and Snyder, "Transaction Log Analyses," 85–89.