

Discovery or Displacement? A Large-Scale Longitudinal Study of the Effect of Discovery Systems on Online Journal Usage

Michael Levine-Clark
University of Denver, michael.levine-clark@du.edu

Jason Price
Statewide California Electronic Library Consortium, jprice@alumni.iu.edu

John McDonald
University of Southern California

Follow this and additional works at: <https://docs.lib.purdue.edu/charleston>



Part of the [Library and Information Science Commons](#)

An indexed, print copy of the Proceedings is also available for purchase at:

<http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

Michael Levine-Clark, Jason Price, and John McDonald, "Discovery or Displacement? A Large-Scale Longitudinal Study of the Effect of Discovery Systems on Online Journal Usage" (2013). *Proceedings of the Charleston Library Conference*.
<http://dx.doi.org/10.5703/1288284315331>

Discovery or Displacement? A Large-Scale Longitudinal Study of the Effect of Discovery Systems on Online Journal Usage

Michael Levine-Clark, Associate Dean for Scholarly Communication and Collections, University of Denver Libraries

Jason Price, Program Manager, Statewide California Electronic Library Consortium (SCELC)

John McDonald, Associate Dean for Collections, University of Southern California

The following is a transcription of a live presentation at the 2013 Charleston Conference. Slides and video are available online at <http://bit.ly/PkzXTv>.

Michael Levine-Clark: Many of us in this room have implemented discovery services in the last five years or so. This is something that academic libraries around the world have been doing as a way of helping our users find information in a way that makes more sense to people who have been raised on Google. When you can search across multiple resources at once, it becomes easier to find information. We have implemented these services and have explicitly featured them in our marketing, or implicitly, at least, by placing a search box front and center on the main page of the library. We have told our users this is where you should start your searches. This is where you should go. We have forced a paradigm shift in how our users find information, and yet we have not looked very carefully at what that shift means in terms of what sorts of information they are actually finding and using. This study attempts to get at that on a large scale across multiple institutions, across multiple sets of publisher data, and across multiple different web scale discovery services.

There have been some local and smaller scale studies that have looked at implementation of a discovery service at a particular institution; looking at usage before and after. Doug Way at Grand Valley State University, for instance, showed that there was an increase in usage in a number of different ways after implementation of Summon at that institution, but it was a single institution and a single discovery service.

Vendor marketing, and we chose WorldCat Local (referring to slide), but we could have easily chosen any number of other sources, shows that at one or two institutions, there is measurable uptake in usage after implementation of a discovery service.

These descriptive statistics show that there is an increase in usage in some ways at some institutions. We want to go deeper than that. We want to go broader than that and see whether there is a significant increase across a range of institutions and a range of services.

We are asking a fairly basic question, “Does implementation of a discovery service impact journal usage?” We know and recognize that the point of discovery services is that they find more than just article content, more than just journal information. They search across books and across a range of other types of information. But we are measuring just journal usage, and within that, we are measuring just publisher-hosted journal content. We are not looking at the increase or decrease of usage in aggregator packages, we are not looking at e-book usage, and we are not looking at print book usage. This is publisher-hosted journal content.

There are four major web scale discovery services: there is EBSCO Discovery Service, or EDS; Ex Libris Primo; Serials Solutions’ Summon; and OCLC’s WorldCat Local. We are looking at all of these different services. Just a quick overview of how web scale discovery services work: they are a single source for finding information from a wide range of different types of information sources. It is a place where you can search for content from books, from articles, locally hosted content, material out on the open web, and so on. They have metadata and full text available, or some combination of those things, and you can search across all of them at once. The content is preindexed or preharvested, depending on how it is added to the system, but the key there is that before you even do your search, there is a database of information that you search for, so it is a fast, single search in a single search box.

The information is pulled in from a variety of different sources. You harvest MARC records from your local OPAC. Publisher metadata and often publisher full text is fed into these beforehand. Abstracting and indexing tools are covered in these as well, (something like MLA Bibliography or PsycInfo). You can pull in information from your institutional repository or any number of other choices of your own, and then things like HathiTrust that are on the open web are also available in these services. If you get excited about the fact that the Red Sox have won three World Series in the last decade you might search on the Red Sox. Yes! And you would find books and articles and newspaper articles and journal articles and so on. It is a wonderful and easy way of searching.

An assumption here is that, at most institutions, you have a relatively stable user base. Unless you are changing your FTE dramatically, adding or subtracting students or adding or subtracting a large number of faculty, the total search effort should stay roughly the same year after year. It is obviously going to fluctuate a little bit, but it is not going to rise or fall dramatically unless the user population or perhaps the curriculum changes in a radical way. But it is unlikely that, at most institutions, that is happening year after year. Put another way, if you have the same number of students, and the same number of assignments, or the same number of faculty writing the same number of articles, there are not more hours in the day for them to do more research, so they are going to be doing roughly the same amount of work year after year after year. If that is true, and you implement a discovery service, and, again, if you implement a discovery service and you say, "This is the place where you should be starting your research," then discovery services should take up an increasing amount of that finite time that any one user has for doing research. Discovery services should draw users away from other search tools, from the abstracting and indexing tools, and from our OPAC, and so on, because we have told our users that you should go here instead of those sources and there, again, is no time to do more searching. Those search tools are more or less effective than discovery services, and so you are going to find different

types of information, and it is going to alter the overall productivity of searches. You are going to find more or less, and it is probably a variation depending on your topic and so on. And it will alter the overall efficiency of users. In other words, they are going to find more or less full text content based on those searches. So we assume then that discovery services have some sort of effect on what people find and how.

There have been prior studies, as I mentioned, and all of these have looked at local use only, so it is a single institution, a single implementation of a discovery service; and they have shown substantial increases after implementation, but this is descriptive data only basically saying, "Yes, it increased." Not why it increased or whether it is statistically significant. Some publishers have reported decreased usage of content, and that is largely anecdotal and certainly not every publisher has reported that. So we are trying to go a little bit deeper in this analysis.

We set out to gather data on libraries that implemented discovery services. We created a list of libraries that had implemented the various discovery services by searching on lib-web-cats and compiling a list of implementations. We then did a survey. We contacted these libraries based on this initial list, and we asked them if, in fact, they had implemented that discovery service that was listed on lib-web-cats. We asked the implementation date, which is crucial; the month; and the year because our entire project is based on before and after usage. We asked if they had the search box on the main page, and we asked if they had done any kind of marketing or PR about this. One hundred forty-nine libraries gave us approval to use their data, and we chose 24 libraries for this initial phase of the study. There are six libraries from each of the four major discovery services, and, again, those are WorldCat Local, EDS, Primo, and Summon.

We tried to do libraries that had implemented at roughly the same time period, and we tried to do libraries that were roughly equal in size across the four discovery services, and these are medium to large academic libraries. We have 20 libraries from the US, and then one each from the UK, from Australia, from New Zealand, and Canada. I think,

importantly, these are all English-language institutions, and we are looking at English-language journal content. There are 10 ARL libraries in this group, and then they range in size down from there. We have got WorldCat book holdings on there as an approximation of library size and the average is 1.1 million volumes. The highest is 2.6 million and then there is a school that has just under 300,000 as the smallest of these institutions.

All of the Primo, Summon, and EDS libraries implemented their discovery service in 2011, and that was the date we were aiming at. We were unable to get WorldCat Local libraries from our initial set that implemented in 2011. There was only one. There were two WorldCat Local libraries in our set that implemented in 2012 and three that implemented in 2010. I am now going to turn it over to John McDonald who is going to talk about the dataset.

John McDonald: My name is John McDonald. I most recently worked at the Claremont Colleges and according to the Brits, I am on “garden leave.” I am about to take a new job at the University of Southern California Libraries as the Associate Dean for Collections. Thank you, Michael, for introducing our study. Now, it is my unenviable task to describe the process we took to collect the actual data, the actual usage data, merge it together, and build a usable dataset so we could do our statistical analysis which would help us answer our research questions using statistics.

Michael already told you about the 24 libraries. It is important to note that, because of the nature of the study, we guaranteed confidentiality to the libraries about the use of their data. We only wanted to report the data in aggregated form, so we are not going to tell you who these 24 libraries are now. At some later date, after we seek approval, and they see the study and what not, I am sure they will allow us to name the libraries, but right now we are just keeping that point confidential. The four discovery tools Michael mentioned, we are not keeping confidential, so everybody knows those four tools. They are the big ones. Then we selected six publishers to use their usage data. These six publishers were nice enough to actually provide us with the libraries’

COUNTER usage data to speed up the processing that we needed to have, so we also guaranteed our publishing study group that we were not going to identify them during this stage of the study. We agreed to let them know if there was anything really very controversial, but we will get to that part a little bit later. But this study does include over 9,000 journals, so you will be able to understand that the publishers we selected were not small publishers that only publish small numbers of journals. We ended up with a dataset of over 9,000 journals, and that included over 150,000 observations. Our definition of the observation is the combination of the library and the journal and its usage. You could take 9,000 journals and multiply it by 24 and come up with a little bit more than 150,000; so we had about 150,000 observations, and it turned out that we had over 140,000 usable observations. Some of the observations had to be removed if we did not have COUNTER reports for all 24 months of the study, and also for a couple of other reasons.

So our methodology: again, it was to take the COUNTER JR1 reports for the 12 months before implementation and the 12 months after the implementation date. The implementation month fell into the “before,” so it could have been implemented on the first day of the month or the thirtieth day the month. We could not get that granular with it but in order to make sure that we were covering the same time period for before and after, we were covering 12 months, so every single school, whether they released it in August in the beginning of the fall semester, or in January of the next semester, each institution would have covered the same set of academic semesters. Our dependent variable is the net change in usage from the 12 months before and 12 months after.

Just a few notes about the dataset, as I just mentioned, we had to exclude any journals that did not have usage reported in a COUNTER report for all 24 months, so if a particular library did not have the title and it was not reported in January of 2011, then we had to leave that journal for that library out of the dataset. A couple of caveats, we had and have limited ability to control for changes in things like aggregator content, as libraries may have selected or deselected aggregator content

that represented publisher content, things like backfile access or expanded holdings that were not represented in a way that we could identify them and remove them from the dataset. And I mentioned earlier about outliers, so there were some outliers in the dataset that we did have to analyze and remove so those could have been journals that had really large increases in usage or really large decreases in uses.

Here is a chart (Figure 1) of the standard scores for every single one of those 159,000 observations, and the data points in red ended up being the outliers, and it was around 100 observations out of 159,000 that we had to remove. Just a point about the statistics: the standard score is the number of standard deviations away from the mean for the whole dataset, and what we selected was anything with 10 or more standard deviations away from the mean, either above or below, were removed from analysis.

We also added up the total journals and developed this pie chart (Figure 2) just to make sure that each of the discovery services were equally represented in the study. As you can see, pretty much every one of them were about equal to about 20 to 30% for each one of them.

This chart (Figure 3) shows all of the institutions grouped by the discovery service, so there are six bars for each of the discovery services from left to right, EDS to WorldCat, and we just wanted to make sure that we did not have large variance across the size of the institutions. So this is the total number of journals that institution had represented in the dataset, and just one interesting thing to note is the larger institutions tend to be able to select Primo and Summon, meanwhile, EDS and WorldCat had a larger range from large to small libraries that had that discovery service.

And then without identifying publishers, we wanted to show the scale of the publishers in the data set and as you can see from this pie chart (Figure 4), we had two very large publishers represented, one midsize publisher and then three smaller more specialized publishers in the

data set but it is six of our biggest publishers in academic research journal literature.

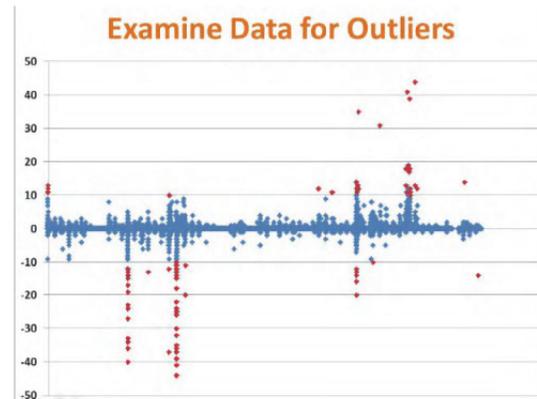


Figure 1. Examine Data for Outliers

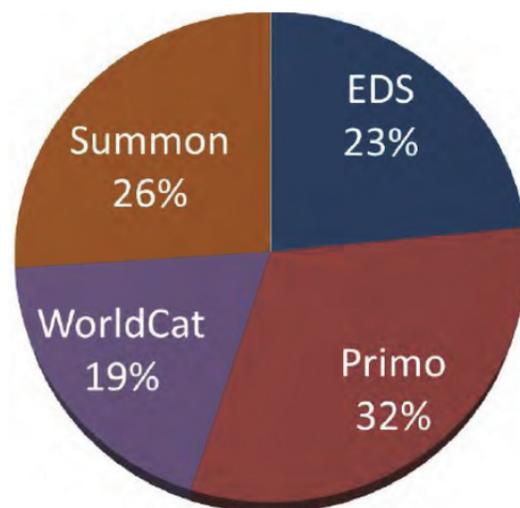


Figure 2. Journal Observations by Discovery Service

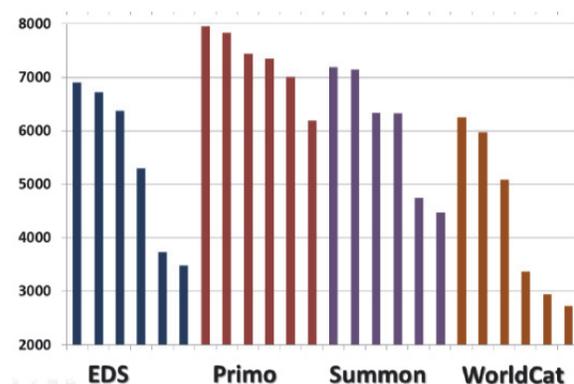


Figure 3. Journal Availability by Institution

And then this bar chart (Figure 5) shows the libraries from left to right by total size or average usage of journals. The blue bars were the year before they implemented their discovery tool, and the red bars are the year after. As you can see, some of the schools went down in their total journal usage for these publishers represented in the dataset, and many others went up; so we just wanted to make sure that there were not any obvious problems with any of institutions in the dataset.

And then in this chart (Figure 6), also another bit of descriptive statistics, we charted all of the publishers grouped by their discovery service and the percentage in change in usage. We wanted to analyze whether any publisher had major increases or major decreases per discovery tool, and as you can see, some of the publishers went way up the journal usage in their average usage and some of them went way down. No single discovery tool increased or decreased the average usage across publishers. So some publishers did very well after implementation of a discovery tool and some did not do as well for certain discovery tools.

So our general observations were that there are large variations for every institution within these discovery tools, major variations by publisher within the discovery tools, and, it seems obvious, but some publishers saw net increases for some discovery tools and not for others, and some experienced actual declines in usage. So now I will turn it over to Jason to describe the actual statistical analysis we performed.

Jason Price: My name is Jason Price. I work for the SCEL in California, and I get to do the fun part which is to try to help us look at differences that appear to be there and ask whether they are real differences or just apparent ones. So that is what inferential statistics do: they say, "Are these means really different? Are the values really different?" The goals of our inferential statistics were really to determine whether those observed differences are significant or resulted from chance effects and to determine which of the three factors: library, publisher discovery service, or even a combination, contributes to determining the differences in usage change at the journal level.

I will start with an exploratory analysis that separates out each factor, but then end with a comprehensive model, and I want to give you a heads up now that the individual factors that we look at were not taken into account as the whole model can be a little misleading, and I will point

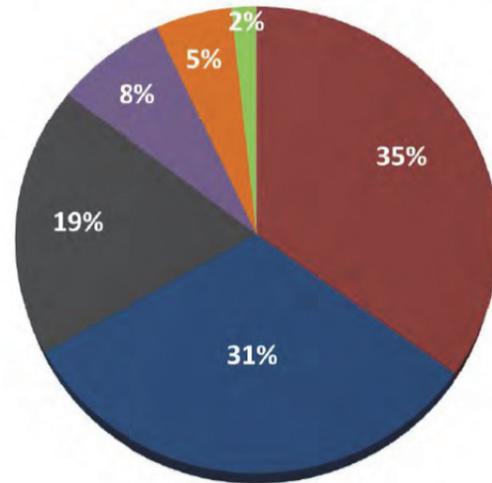


Figure 4. Journal Observations by Publisher

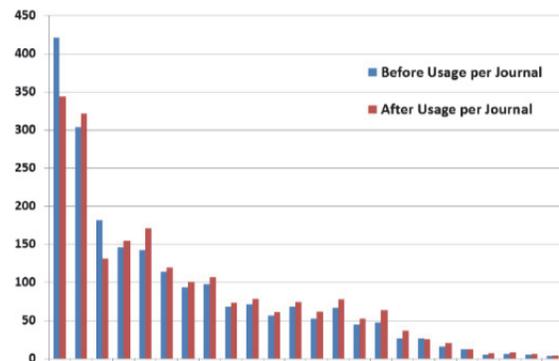


Figure 5. Average Journal Usage by Library

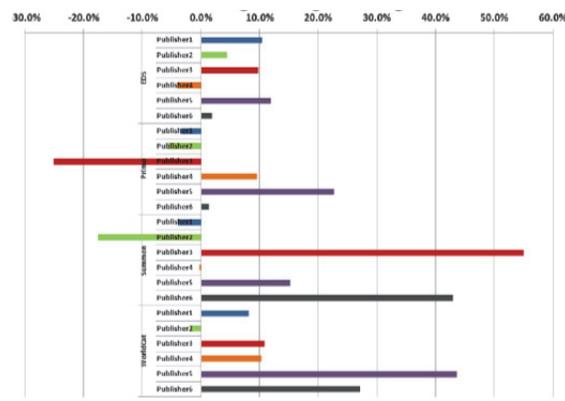


Figure 6. Percentage Change in Usage

that out along the way and show you the complete model at the end.

If you will bear with me, some basic statistical information: what is ANOVA? It is the analysis of variance, and what it does is it looks at each individual observation that we take, each individual journal that we look at, its change from before and after, and it says how much of that is due to things that we are measuring and know about, which library did it come from? Which publisher published the journal? Which discovery tool was used and impacted that change? And it separates those factors to fit from the residual, which is the error or the things that we cannot explain, with the factors that we measured. So for a particular observation, let us just say it is a change of positive 17 downloads, we might know that a library, on average, has two more downloads than the average of all libraries, and a particular publisher on average has three less downloads than the average publisher, and the discovery service impact might be 10 more than the average discovery tool result. Those things are the things that we can take into account and ask about what their influences is. But if our residual error, that part of the observation that we cannot explain, is high relative to the numbers we get from looking at the factors we are measuring, then we have to say, "Well, we cannot really tell." These things are there but the error swamps out any effect that we might be able to see.

Put another way, what we are doing is we are testing to see whether the means for the levels like the different libraries within a factor are distinguishable from each other. We have, and I will show you, F Ratio tests. These are numbers that look at the average variability due to the factor versus the average variability due to chance error. If the chance error is high relative to the average variability of the factor, then our F Value will be close to one and we will not see difference. We will not be able to distinguish them. If our average variability factor is large relative to the chance error, then F Value will be higher and we will be able to detect a difference. So that is it for the statistical background. Now, let us get to the questions.

In response to our first question, "Does usage change vary across libraries?" (Figure 7), we see the 24 libraries sorted by mean change across the x axis and the mean change plus or minus two standard errors on the y axis, so standard errors, these bars that you see, are a measure of the variability around that mean. In general, when the bars overlap, then between different levels of the factor we cannot say they are different; when they do not overlap they are different. Our initial observation here would be that there are some differences among libraries in their mean change. The overall average change was 8.5 uses on the journal level, and our F Ratio of about 32 tells us that institution alone is a significant predictor of the mean change in usage after discovery implementation. Whenever we look at the significance value, if it is below 0.05 that indicates that we can detect a difference. If it is above 0.05 then we cannot. But, for single factor ANOVA's, single looking at just one variable, we are ignoring the impact of the different discovery tools and journal publishers on this mean change, so this tells us something, but it does not take all the information we have into account.

A couple more, just to give you that feel for the individual factors: now we ask whether usage change varies across publishers (Figure 8), and the grand mean for the change in usage here is about 8.9. One publisher appears to have a mean change that is not significantly different from zero, the one on the left. The one on the right seems to have a mean change that is significantly higher than the other four, and then the other four are right along that mean. Again, our significant F Value and nonoverlapping error bars suggest that mean change did vary across

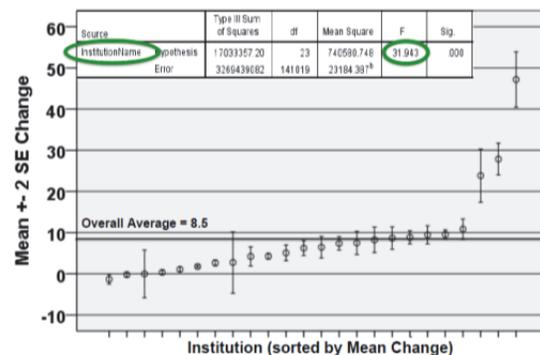


Figure 7. Does Usage Change Vary Across Libraries?

publishers, but this is one of those cases where when we just look at this factor, we actually think that there are differences among publishers, that a publisher is significant, but when we put everything together, I will show you at the end, that publisher drop out is an important factor at least looking at it just alone. It is, in fact, not relative to other factors that we take into account.

So this is the big slide (Figure 9). The data shows that the mean usage increase was positive for all discovery tools. So in each case, the usage from the year before to the year after increased, although we cannot distinguish these from industry-wide increases, such as the fact that usage tends to go up year to year or because we did not measure a set of libraries that did not implement discovery tools. Whether they are significant from the background, we do not know, but we can say they are different from each other, so it appears that Primo and Summon had a greater increase in the usage of the content we studied than did EDS or WorldCat Local.

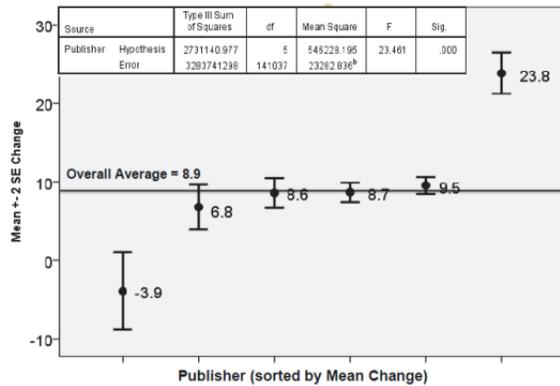


Figure 8. Does Usage Change Vary Across Publishers?

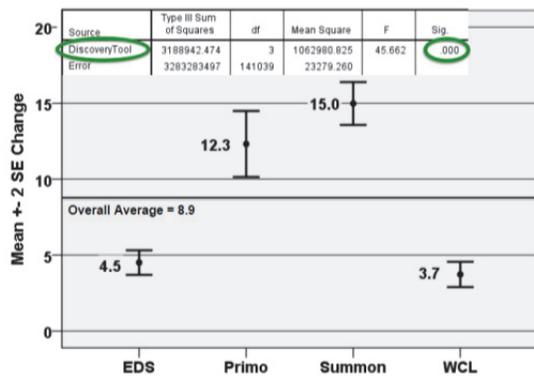


Figure 9. Does Usage Change Vary Across Discovery Service?

But when we start to make the model more complex, when we start to look at discovery service and publisher together, we look at a more complex relationship (Figure 10). The means that we just looked at those four different places that we saw the mean change; now we see some of the background behind that, to see which publishers had more mean change or less. What our model does, is it allows us to look at these things at the same time; so in addition to asking whether we can detect a difference across discovery service and publisher, this two-way model (Figure 11) addresses whether the impact of discovery is equivalent for each publisher, and, in fact, this is one of the reasons that we wanted to do the study in the first place, is that we had heard rumors that some discovery tools were differentially affecting publishers to the point that there was concern among publishers who were finding their usage was falling. So what the interaction between these two things does is it basically asks statistically, "Are these lines parallel or not?"

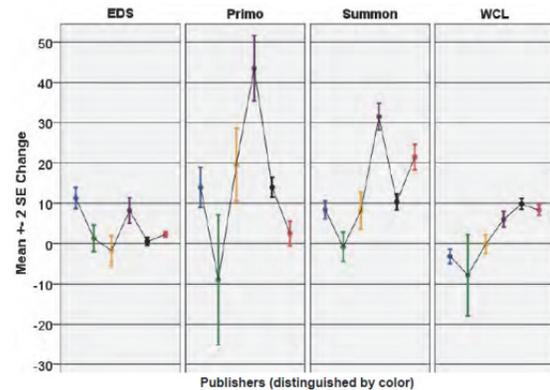


Figure 10. Does the Effect of Discovery System Vary Across Publishers?

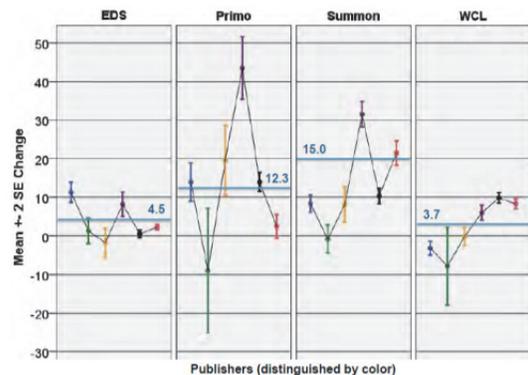


Figure 11. Do the Discovery Service Means Differ in the 2-Way Model?

So, what we have done is broken down those bars by each discovery tool across the panels, and the different colors represent the different publishers' content, and so again we are asking, "Are these lines parallel?" So first we can look at our factor based on discovery tools. We see those means and can ask, "Are those means significant when we look at both?" We can look at the means across the publisher, the mean increase for the publisher, and ask if those are significant, and then we could look at that combination. Our statistics now bring these things together, and what we notice right off is that when we look at publisher we do not find a significant effect of publisher in the combined dataset, so those kind of dropout of meaningfulness in this model. Then, discovery tool similarly is above 0.05, and so even the effect of discovery tool alone disappears when we just look at these two factors. However, what we do see is that the interaction of discovery tool between publishers is significant, so we are finding that discovery tools affect different usage of different publishers' content differently.

But to really do this right, we have to think about the institutions, too. Some of you may be wondering, "Well, what about how much variance there is under the institutions?" So the full model, this is the whole thing put together, the statistically most rigorous way to approach the data, we need to recognize that we looked at the four different discovery tools and each publishers' content was represented for each of those four discovery tools. However, each library only used one discovery tool, right? So libraries are actually nested within a discovery tool. Library 1 only used Discovery Tool 1. Library 7 only used Discovery Tool 2. So the beauty of the statistics is that it can partition that variance and help us to look at the whole model and figure out what is going on, and so that is what we do here. We find that, again, publisher alone is not a significant predictor of the mean change in usage. Discovery tool and institution nested within discovery tool are still significant predictors, so we do notice differences just among those alone, but then the impact of the interaction between discovery tool and publisher, that differential effect and of the impact of different publishers' content within institutions are also significant.

So in summary, our results ask, "Can we detect differences between discovery tools, services, publishers, and libraries and/or their interactions?" Discovery service, we saw an effect; publisher: no effect; library: yes. There is variation due to library, and, furthermore, there is a differential discovery of service effect by publisher and a differential library effect by publisher. One of the things we want to do going forward is to make sure this model is giving us the best answer that it can, that we are bringing in other libraries, and we are addressing other possible parts or impacts, influences on these changes. I am going to turn it over to Michael.

Michael Levine-Clark: So we will wrap up here. One of the things we discovered, but we knew this already, is that analyzing usage is a complex task. It is not as easy as just saying yes or no, it increased or decreased. No discovery service increased or decreased usage across all libraries or all publishers. There were variations across the board. Discovery service on its own was a significant predictor of usage change, and the interaction of the discovery service and the publisher was significant as well. We know that we need to control for libraries that do not have a discovery service, and we know that we need to control for the size of the institution. So next steps, what we need to do going forward, because this is the first phase of this study: we need to take into account aggregator full text availability as a factor. We need to control for institution size or enrollment profile, are there variations by type of institution that may not have a discovery service at all? For publisher size, for journal subject, we need to look at overall usage trends across libraries in general, so a control group that did not implement a discovery service during this time period, and then we also need to look at things like configuration options in the discovery service. How were they implemented, how does that interplay with choice of link resolver, and so on? We will be doing a follow-up presentation at UKSG in April, and we will have a control group then and bring in the additional libraries and bring in these other variables and this further analysis as we have noted above. Now we are going to sit and take questions.