

2018

## Comparing Apples and Bananas? A/B Testing for Discovery System Optimization

Sarah Dahlen

*California State University, Monterey Bay, sdahlen@csumb.edu*

Kenny Garcia

*California State University, Monterey Bay, kengarcia@csumb.edu*

Kathlene Hanson

*California State University, Monterey Bay, khanson@csumb.edu*

Follow this and additional works at: [https://digitalcommons.csumb.edu/lib\\_fac](https://digitalcommons.csumb.edu/lib_fac)

---

### Recommended Citation

Dahlen, S., Garcia, K., & Hanson, K. (2018). Comparing apples and bananas? A/B testing for discovery layer optimization. Proceedings from CARL 2018: The Academic Library in Times of Change. <http://conf2018.carl-acrl.org/conference-proceedings/>

This Conference Proceeding is brought to you for free and open access by the Library at Digital Commons @ CSUMB. It has been accepted for inclusion in Library Faculty Publications and Presentations by an authorized administrator of Digital Commons @ CSUMB. For more information, please contact [digitalcommons@csumb.edu](mailto:digitalcommons@csumb.edu).

# Comparing Apples and Bananas? A/B Testing for Discovery System Optimization

Presented at the California Academic & Research Libraries 2018 Conference

April 13-15, 2018

Redwood City, California

Sarah Dahlen, Research & Instruction Librarian, California State University, Monterey Bay

Kenny Garcia, Research & Instruction Librarian, California State University, Monterey Bay

Kathlene Hanson, Electronic Resources Coordinator, California State University, Monterey Bay

## **Abstract**

Discovery systems have changed the ways in which students are able to search academic library resources by providing a simplified, customizable user interface. Librarians often make discovery system customization decisions based on instincts about what will work best for students or on small usability studies. A/B testing, commonly used by commercial enterprise but infrequently by libraries, employs two simultaneous, live versions of a web interface to gauge the effects of changing variables. This method has the advantage of reaching a large number of users performing authentic search tasks. The authors combined A/B testing with scenario-based usability testing to explore variables such as facet labels, facet order, and placement of search options. The results provide insight into the extent to which interface changes prompt users to employ available search options, allowing libraries to make decisions about discovery system customization driven by user data.

## **Background**

Along with the other 22 California State University libraries, the CSUMB Library recently migrated to Ex Libris' Primo as our discovery system, which went live in June 2017. As libraries adopt discovery systems and work to optimize their performance, we need to attend to the user experience when making configuration and implementation decisions. Librarians spend a lot of time customizing user interfaces based on what we assume to be best for our users, but usability studies can provide user data to better drive these decisions. The specifics of our testing configuration will be of particular interest to libraries using Primo, but the lessons learned about A/B testing and what our data tell us about user behavior may be useful to libraries using other discovery systems as well.

## **Discovery Usability Testing in Literature**

Usability testing has become the norm for validating ease of use and effectiveness of library websites (Prommann & Zao, 2015) and is becoming prevalent as a way to inform configuration of library discovery systems. Case studies mainly discuss the use of scenario-based usability testing and gathering of information either through think-aloud methods or post-scenario interviews (Brett, Lierman, & Turner, 2016; Perrin, et al., 2014). Various types of think-aloud protocols exist, from those with little to no intervention to those with active intervention (Alhadreti & Mayhew, 2017). Ultimately, the goal of all these is to glean information about how users complete tasks. Findings related to discovery systems include the infrequent use of facets while searching, with users preferring to modify their search terms instead (Brett, Lierman, & Turner, 2016).

Other types of usability testing are beginning to be used by academic libraries, for example, hierarchical task design (Prommann & Zao, 2015) and event tracking (Hanrath & Kottman, 2015), though still infrequently. In addition, though prevalent in the commercial sector (Fichter & Wisniewski, 2017), A/B testing in academic libraries is not discussed in peer-reviewed literature. The addition of A/B testing to libraries' usability toolbox can lead to useful information from authentic searches without direct user intervention (Fichter & Wisniewski, 2017).

## **Methodology**

We used a mixed methods approach to our usability testing in order to benefit from the complementary insights provided by qualitative and quantitative methods. Our qualitative data came from scenario-based usability testing in which student participants were given a scenario that prompted them to complete a task using Primo (see Appendix 1). As students completed the task, they were encouraged to talk aloud about what they were thinking or seeing. This narration, in addition to post-scenario interviews, provided insights into the ways that students search and the challenges they encounter. Following Nielsen (2000), our sample size for each of our four sets of scenarios was five participants, for a total of 20 participants.

Quantitative data was collected using A/B testing, a method in which there are two simultaneously live versions of a website (or in our case, of Primo), that are identical with the exception of one change. This allows us to see the effect of that change on user behavior with all other variables, including the important variable of time, being controlled. User traffic was automatically directed to one of the two versions of Primo (version A and version B), creating a large sample size of users engaged in authentic information-seeking tasks.

In A/B testing, user behavior is analyzed using conversion rates for each version of the website. The difference in conversion rates, or percent of users performing an action of interest, is compared for the two versions. We used a Z-test for two proportions to determine whether the difference in conversion rates was statistically significant (see Appendix 2).

Services such as the fee-based Optimizely and the free Google Optimize can streamline A/B testing. In our case, testing a discovery system outside of our web domain prevented us from

using one of these services and led us to a do-it-yourself approach in which the A/B tests were set up with two alternate views in Primo. The alternate views were randomly assigned through a script embedded in the single search box on the library home page. To make the tests authentic, users were not notified about alternative views.

While we tested changes such as facet order, number of items to display per facet, the Brief Results page, and the order of sections in full display view (which required a CSS change), we were not able to test everything of interest, most notably label changes. Changes to labels require a change to mapping and code tables, which would have made changes across all of the views. For example, if we wanted to revise the label for the “Personalize” button, the change would appear on both alternate views and the main view. We did receive feedback on potential label changes from students through post-scenario interviews.

Primo Analytics was used to gather data for the A/B tests. We created Analysis modules to run reports, and added filters in the analysis criteria to filter out views, dates, facets selected, and actions taken. Working with a new system required a lot of testing to see what the reports would produce. In some cases, what was being reported was unclear, so we created a test view to look at analytics. For example, we were unsure if clicking on the facet label to expand the list of results counted as an action or if Primo Analytics just counted the action of clicking on a facet item. The development of additional documentation was and will be helpful in the future. We reviewed the Orbis Cascade Primo toolkit (Orbis Cascade Alliance, 2018), asked the CSU ULMS committee and campus project managers questions, and documented our own local processes.

## **Results**

Much of our usability testing focused on facets, in particular the format and the subject facets (Ex Libris labels the latter “topic” by default). For the format facet, we used an A/B test to determine whether the sort order of the formats (listed either alphabetically or by frequency of occurrence) had any effect on how often the books format was chosen (see Figure 1). We focused on books because we anticipated that some searches would have so few book results that “books” would not appear in the top seven formats that are displayed when sorted by frequency. In fact, “books” generally does appear in the first seven formats, and the test showed that listing format options alphabetically or in terms of frequency had no effect on how often the book format was chosen.



Figure 1: Format facet in frequency and alphabetical order

Several of our scenarios provided insights into use of the format facet as well. Various scenarios asked students to find an encyclopedia, an ebook, and articles without specifying using the facet. None of the students used the format facet “reference entries” to limit to encyclopedias (see Figure 2). When asked, the students said “reference entries” meant nothing to them and when shown an example and asked to re-label, a few suggested “encyclopedias.”

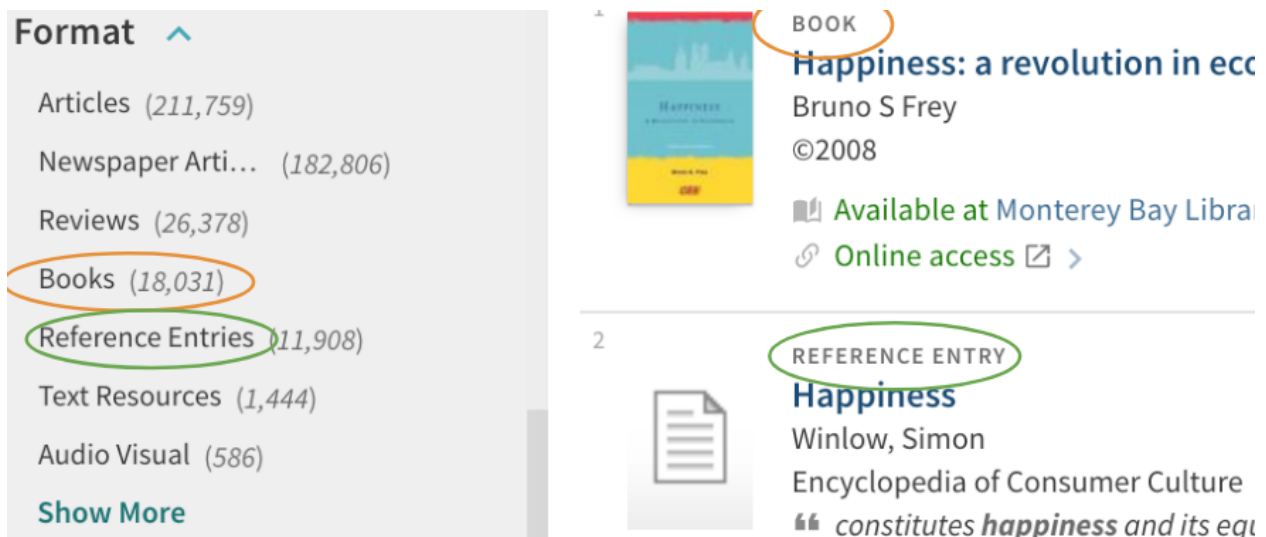


Figure 2: Reference Entries facet item

When prompted by a scenario to find an ebook, students unsuccessfully looked for ebooks among the facets. Our current configuration does not list ebooks as a separate facet item from books. Most students were able to find an ebook simply by scrolling through results until

they identified a result with “online access.” A few went into individual records to look for an indication that a book was available online. One used the “full-text online” limiter (see Figure 3).

The screenshot shows a library search interface. On the left, there are two filter panels. The top panel, titled "Available at CSUMB", has three items: "Full-text online (30,812)" (circled in orange), "Available in the Libr... (463)", and "Peer-reviewed Jour... (366)". The bottom panel, titled "Format", has five items: "Books (18,031)" (circled in orange), "Reference Entries (11,908)", "Audio Visual (586)", and "Government Docume... (80)". On the right, there are two search results. The first is a "BOOK" titled "Happiness: a revolution in eco" by Bruno S Frey, published in 2008. It has a book cover icon and a link to "Available at Monterey Bay Librar" with a circled "Online access" link. The second result is a "REFERENCE ENTRY" titled "Happiness" by Winlow, Simon, from the "Encyclopedia of Consumer Culture". It includes a quote: "constitutes happiness and its equ self Happiness in... Happiness".

Figure 3: Finding ebooks

During scenarios where students were asked to find articles, some successfully used the “articles” or “newspaper articles” facet items (see Figure 4). Others simply scrolled through the results and selected those that were labeled as “article” or “newspaper article” in the brief results display.

The screenshot shows a library search interface with filters and search results. On the left, there are two filter panels. The top panel, titled "Active filters", has one item: "Articles X" and a "Reset filters" button. The bottom panel, titled "Refine by", has one item: "Search Beyond CSU" (unchecked). The middle panel, titled "Sort by", has one item: "Relevance". The bottom panel, titled "Available at CSUMB", has one item: "Peer-reviewed Jour...". On the right, there are two search results. The first is a "NEWSPAPER ARTICLE" titled "Upward Mobility: A Better Direc" by Riley, Jason, from the "Wall Street Journal, Jun 28, 2017, p.A.". It includes a quote: "dear to the Black Lives Matter mov movement got its... Will Black Lives M" and a link to "Full text available". The second result is a "NEWSPAPER ARTICLE" titled "Unions and 'Black Lives Matter'" by Johnson, Stephon, from the "New York Amsterdam News, Aug 6-Aug". It includes a quote: "to align themselves with Black Live lives Matter... since its inception, the 'I".

Figure 4: Articles and newspaper articles facet items

For the subject facet, we explored the effects of its location on the results page, whether it was collapsed or expanded, and its label. A/B tests were performed to ascertain whether location had an effect on use of the subject facet and whether this differed when collapsed or expanded (see Figures 5 & 6). When collapsed, there was no significant difference in use regardless of whether the facet was located higher or lower in the facet list. However, when the facet was expanded there was a significant increase in use when the facet appeared higher up among the facets. Additionally, with the location held constant, there was more use of the subject facet when it was expanded than when it was collapsed.

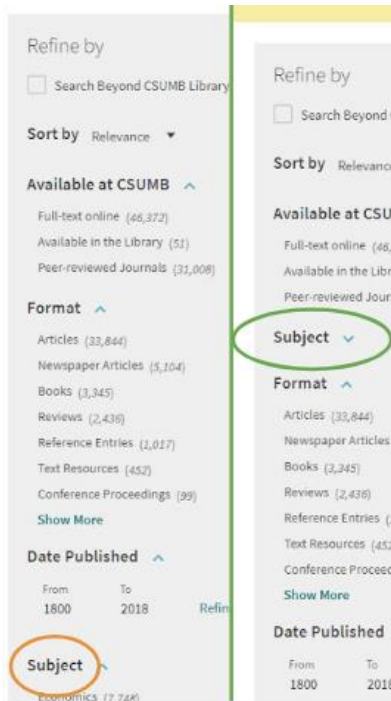


Figure 5: Collapsed subject facet

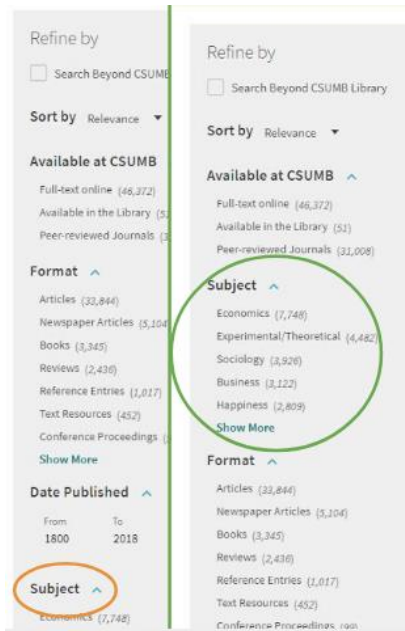


Figure 6: Expanded subject facet

During post-scenario interviews, we showed students the collapsed subject facet and asked what “subject” meant in this context. Most students imagined that this option would give them disciplines, and were surprised to see it was more specific than that (see Figure 7). A few students noticed the differing levels of specificity of the subjects. When asked about an alternative label, a few students suggested “topic,” though most did not have a suggestion. The confusion around “subject” and “topic” became particularly apparent when a few students used the advanced search screen to conduct their searches. In advanced search, several students misused the subject search from the dropdown, seeing it as a topic rather than an official subject heading (see Figure 8). It is clear that students are unsure of what these terms mean and likely do not understand what controlled vocabulary searching is when compared with keyword searching.

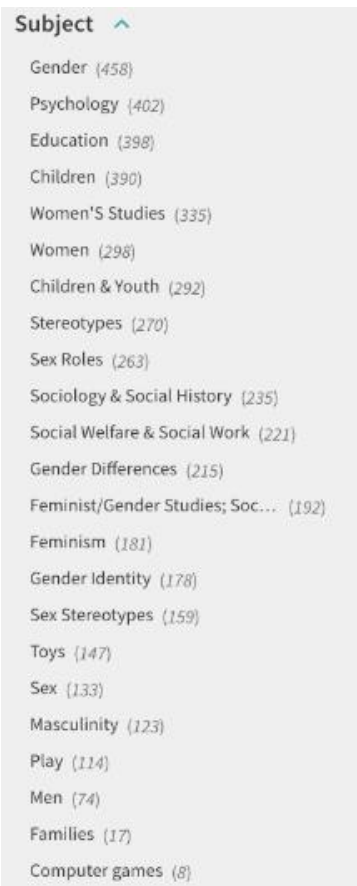


Figure 7: List of subjects/topics

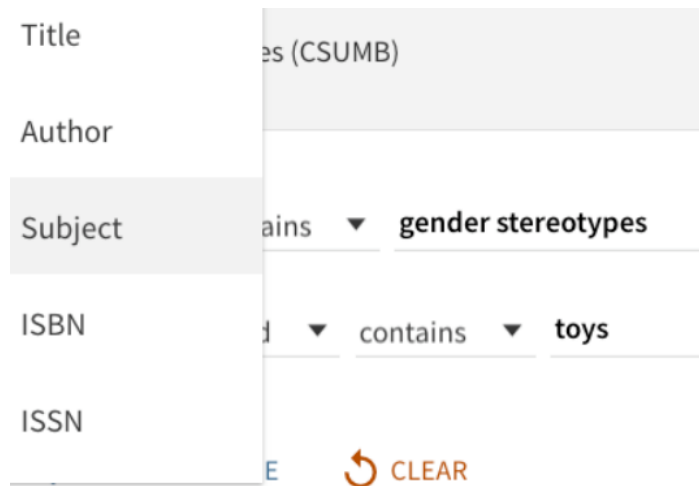


Figure 8: Subject in advanced search

The Primo results page includes a “search beyond” option (see Figure 9) that expands the search to the entire Primo Central Index, rather than just the articles to which our library has full-text access. When asked what they imagined this option would do, most students understood that it would expand their search, but seemed to think about this in only terms of physical items, such as books, rather than electronic articles. It appears that the current label is not clearly communicating to students the function of this option.



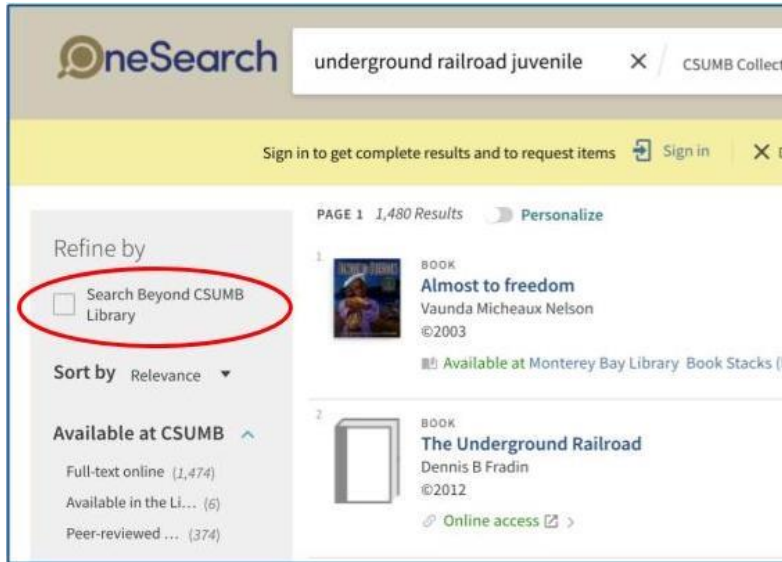


Figure 9: Search Beyond the CSUMB Library

At the end of the results page is a link to “load more results” (see Figure 10), adding items beyond the initial 10 that we had set as our default. In an A/B test where the variable was the default number of results, we found that users are more likely to click “load more results” when the initial results list includes 20 results than they are when it includes 10. This may seem counterintuitive, but a possible explanation is that when fewer results are visible, users may not see anything they like and revise the search. When more results are visible, users may see more promising results in the expanded range and continue loading additional results. When asked whether they usually look through more than one page of results, student answers were mixed. Of those who said that they do not, they assume that the best results are probably at the beginning, and would rather revise their search or choose a different database than load more pages.

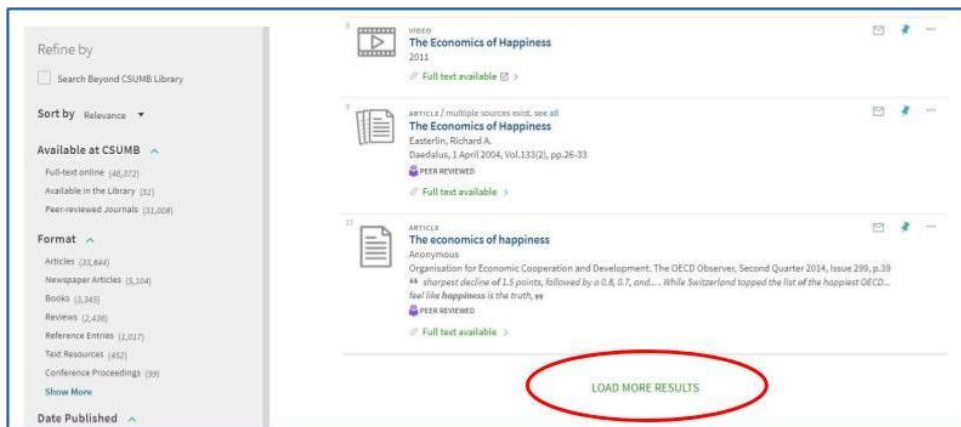


Figure 10: Load More Results

Primo's results page includes the option to "personalize" a search (see Figure 11). When selected, the user can choose disciplines to prioritize among their results (see Figure 12). We had two different scenarios, completed by different sets of students, that involved narrowing a search by discipline. In neither case did any of the students use the personalize option to complete this task. In post-scenario interviews, we had students look at the unexpanded personalize button (see Figure 11) and asked what they imagined it would do. Most students guessed that it had something to do with tracking their past search history and suggesting similar results. They were surprised to see what it actually did, and some seemed to think that it was useful. "Subject areas" or "disciplines" were suggested as alternate labels.

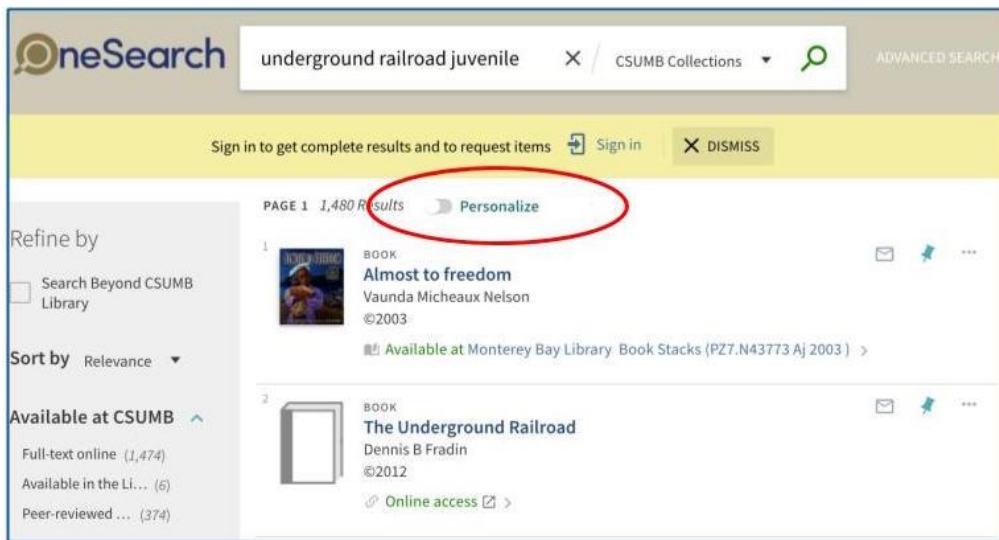


Figure 11: Personalize

### Personalize the results

You can help us provide you with better result matches by focusing the search on your preferred disciplines. Please select up to 5 disciplines.

- |  |  |
|--|--|
| <input type="checkbox"/> Agriculture & Forestry                | <input type="checkbox"/> Nursing                     |
| <input type="checkbox"/> Arts & Humanities                     | <input type="checkbox"/> Physics                     |
| <input type="checkbox"/> History & Archaeology                 | <input type="checkbox"/> Psychology                  |
| <input type="checkbox"/> Languages & Literature                | <input type="checkbox"/> Sciences                    |
| <input type="checkbox"/> Philosophy & Religion                 | <input type="checkbox"/> Biology                     |
| <input type="checkbox"/> Business & Economics                  | <input type="checkbox"/> Chemistry                   |
| <input type="checkbox"/> Engineering                           | <input type="checkbox"/> Earth Sciences              |
| <input type="checkbox"/> Computer Science                      | <input type="checkbox"/> Mathematics                 |
| <input type="checkbox"/> Materials Science                     | <input type="checkbox"/> Social Sciences             |
| <input type="checkbox"/> Law                                   | <input type="checkbox"/> Anthropology                |
| <input type="checkbox"/> Library & Information Science         | <input type="checkbox"/> Education                   |
| <input type="checkbox"/> Medicine                              | <input type="checkbox"/> Geography                   |
| <input type="checkbox"/> Diet & Clinical Nutrition             | <input type="checkbox"/> Journalism & Communications |
| <input type="checkbox"/> Pharmacy, Therapeutics & Pharmacology | <input type="checkbox"/> Political Sciences          |
|  | <input type="checkbox"/> Public Health               |
|  | <input type="checkbox"/> Sociology                   |

Figure 12: Personalize expanded

The item record is another area where Primo allows some configuration, and we conducted an A/B test gauging the effect of the placement of certain components of the record. The “send to” bar, containing links for email, permalink, printing, and bibliographic management tools, generated many more clicks when it was at the top of the item record page than at the bottom, below the “fold” (see Figure 13). While such a dramatic change in page location had a statistically significant effect on user behavior, a smaller change in location did not. The difference in location of the “sign in” bar in Figure 13 (first versus second item on the page) did not have a significant effect on user sign ins. While location matters, it seems to be a matter of degree.

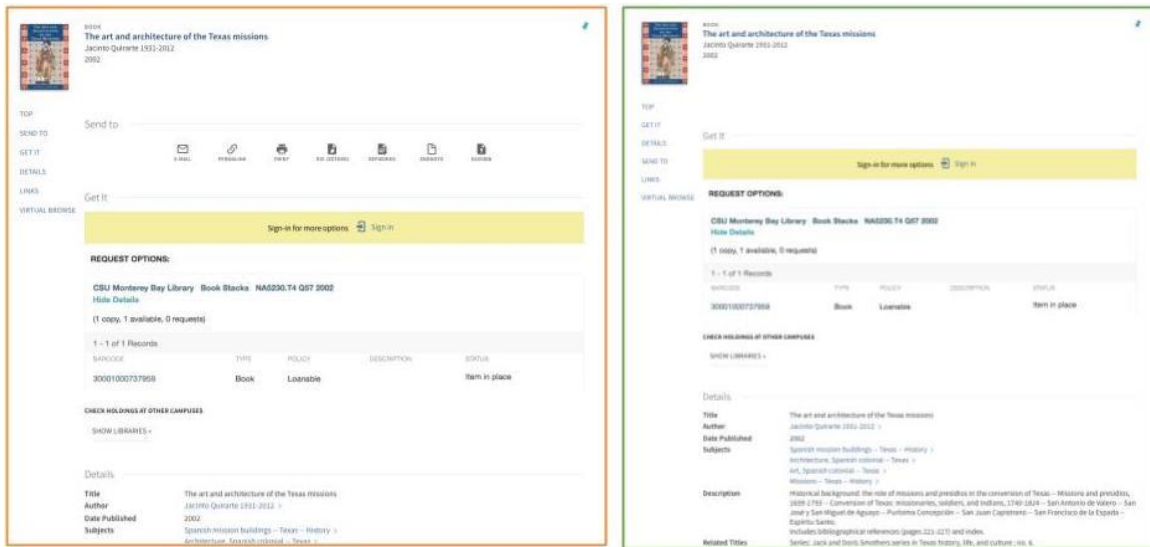


Figure 13: Item record

One of our scopes in Primo allows searching for books and media at all CSU campuses (see Figure 14). CSU+ is the system for requesting books from other campuses, but the link to it appears only when the user is signed in. In one scenario, students were asked to find and request a book from another CSU campus. While many of them found the appropriate scope, they struggled to find the link to CSU+ because they were not reading the text indicating that they needed to sign in (see Figure 15). Simplified text might help alleviate this problem.

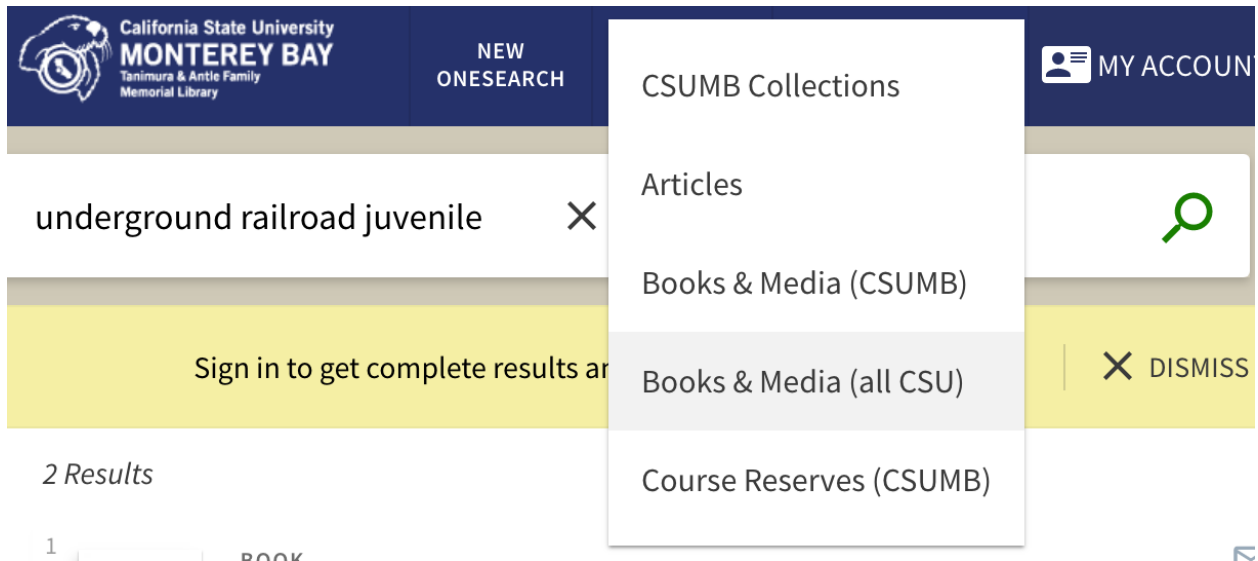


Figure 14: Scope to search all CSU libraries

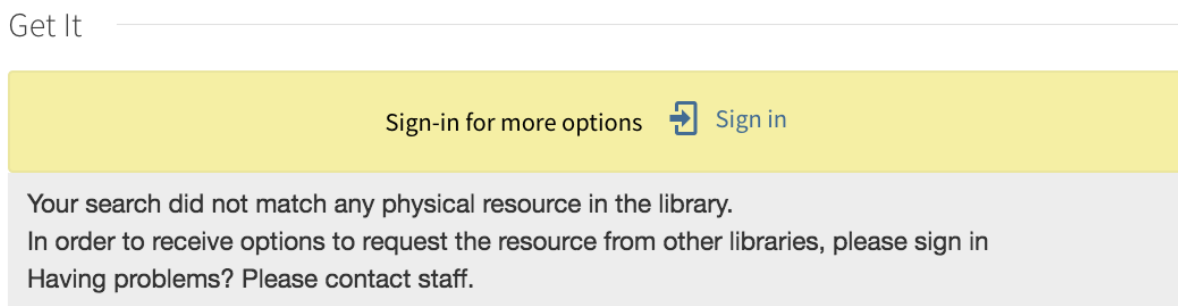


Figure 15: Sign in prompt to access CSU+

It is worth noting that there were some tasks in Primo that students completed without difficulty. These included finding known items, limiting by date, and saving item records, though their approach to the latter varied among using the email, permalink, and save options.

## Discussion

Beyond enabling us to make informed changes locally, one of the reasons our usability findings are worth sharing is that previous research we conducted indicated that discovery system configuration has implications not just for usability but for the quality of information that students select. It has been well established that students tend to prefer discovery systems over traditional database interfaces (Asher, Duke, & Wilson, 2013; Gross & Sheridan, 2011; Lundrigan, Manuel, & Yan, 2015; Rose-Wiles & Hofmann, 2013), and that has held true on our campus as well. During a previous research project, students ranked Summon (our discovery

system at that time) as easier to learn, more useful for coursework, and producing better search results when compared to a traditional database, Social Sciences Abstracts (Dahlen & Hanson, 2017).

In addition to comparing a discovery system with a traditional database, our previous research included two configurations of Summon: a default version that was inclusive of formats and subjects and a pre-scoped version that was limited to social science disciplines and excluded newspapers (though it included other non-scholarly formats). Students conducted searches on a specified topic in each configuration of Summon and in Social Sciences Abstracts. They were asked to find two of the “best quality” articles from each search tool, with that term left to their own definition. The authority of the articles chosen was scored using a taxonomy developed by Leeder, Marker, and Yakel (2012). The articles our students chose from pre-scoped Summon and Social Sciences Abstracts had greater authority than those from default Summon, and this difference was statistically significant. The implication is that the configuration of discovery systems matters, not only to students’ ability to complete tasks, but also to the quality of information they choose.

It does not make good sense, however, to have the library’s main instance of our discovery system pre-scoped by subject or format. Because of this, it is important to make sure that students can easily modify their search in ways that will guide them toward more authoritative results.

## **Conclusions**

Based on our findings, the changes to our Primo interface we are considering include: leaving the subject facet expanded regardless of its location; moving the subject facet higher in the list of facets; changing labels such as “Personalize”, “Search beyond CSUMB Library”, and “Reference Entries”; revising the text for CSU+; leaving “Send To” at the top of the item record; and expanding the brief results list to display 20 results.

For libraries contemplating whether A/B testing is the right tool for their purposes, our recommendation is to consider several factors. The large amount of data we gathered from authentic searches was incredibly valuable in informing decisions. However, our process was time and labor intensive, and there were limitations on what we could test. We only have access to one instance of Primo, which does not allow for testing label changes on individual views. We had the advantages of staff with the technical skills to write the script randomizing the alternate views and the advanced administrative privileges to make changes to the library’s website. A/B testing using a tool like Optimizely or Google Optimize would cut down significantly on the workload and could work well for usability testing for library websites.

Using mixed methods for usability testing, particularly combining user data from A/B tests with scenario-based testing and post-scenario interviews, can help us learn more about how students are using Primo and lead to informed decisions about changes to the Primo interface.

Discovery system customization can be most effective when based on librarian expertise alongside user data.

## Works cited

- Alhadreti, O., & Mayhew, P. (2017). To Intervene or Not to Intervene: An Investigation of Three Think-Aloud Protocols in Usability Testing. *Journal of Usability Studies*, 12(3), 111–132.
- Asher, A. D., Duke, L. M., & Wilson, S. (2013). Paths of discovery: Comparing the search effectiveness of EBSCO Discovery Service, Summon, Google Scholar, and conventional library resources. *College & Research Libraries*, 74(5), 464–488.
- Brett, K., Lierman, A., & Turner, C. (2016). Lessons learned: A Primo usability study. *Information Technology & Libraries*, 35(1), 7-25. doi:10.6017/ital.v35i1.8965
- Dahlen, S. P. C., & Hanson, K. (2017). Preference vs. authority: A comparison of student searching in a subject-specific indexing and abstracting database and a customized discovery layer. *College & Research Libraries*, 78(7), 878–897.
- Fichter, D., & Wisniewski, J. (2017). Boost your website performance with A/B testing. *Online Searcher*, 41(2), 66-68.
- Gross, J., & Sheridan, L. (2011). Web scale discovery: The user experience. *New Library World*, 112(5/6), 236–247. <http://dx.doi.org/10.1108/03074801111136275>
- Hanrath, S., & Kottman, M. (2015). Use and usability of a discovery tool in an academic library. *Journal of Web Librarianship*, 9(1), 1-21.
- Leeder, C., Markey, K., & Yakel, E. (2012). A faceted taxonomy for rating student bibliographies in an online information literacy game. *College & Research Libraries*, 73(2), 115–133.
- Lundrigan, C., Manuel, K., & Yan, M. (2015). “Pretty rad”: Explorations in user satisfaction with a discovery layer at Ryerson University. *College & Research Libraries*, 76(1), 43–62. <https://doi.org/10.5860/crl.76.1.43>
- Nielsen, J. (2010). Why you only need to test with 5 users. Retrieved April 13, 2018, from <https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>
- Orbis Cascade Alliance. (2018). *Primo Toolkit*. Retrieved from <https://www.orbiscascade.org/primo-toolkit>.
- Perrin, J. M., Clark, M., De-Leon, E., & Edgar, L. (2014). Usability testing for greater impact: A

- Primo case study. *Information Technology & Libraries*, 33(4), 57-66.
- Prommann, M., & Tao, Z. (2015). Applying hierarchical task analysis method to discovery layer evaluation. *Information Technology & Libraries*, 34(1), 77-105.
- Rose-Wiles, L. M., & Hofmann, M. A. (2013). Still desperately seeking citations: Undergraduate research in the age of web-scale discovery. *Journal of Library Administration*, 53(2-3), 147-166. <https://doi.org/10.1080/01930826.2013.853493>

## Appendix 1: Usability Scenarios

|   | <b>Task</b>                      | <b>Scenario</b>  |
|---|----------------------------------|--|
| 1 | Find known book                  | Your instructor has told you that the library might have one of the required texts for your course, “History of Anthropology” by Thomas Eriksen. Figure out whether the library has the book and how you can check it out or read it.  |
| 2 | Find books by topic              | Your assignment requires that you use a book as one of your sources for your paper on the history of women’s rights in Middle Eastern countries. Use the CSUMB Library’s website to find a book that you might use.  |
| 3 | Find scholarly articles by topic | You are writing a research paper on the migration patterns of sharks and your instructor has asked that you find scholarly articles as sources. Use the CSUMB Library’s OneSearch to find one article that you might use.  |
| 4 | Save item for later              | You’ve found a book in OneSearch that you’re considering using for an upcoming paper for one of your classes. You don’t have time to get it from the shelf right now, and you left your smartphone at home. Using OneSearch, find a way to come back to the book record later. |
| 5 | Narrow search by discipline      | You are writing a paper about gender stereotypes in children’s toys, and your assignment requires that the sources for your paper be from a sociological perspective. Find one book or article that you think you could use.   |
| 6 | Narrow search by format          | You are writing a research paper on how the media covers social justice activism. Find four newspaper articles on the Black Lives Matter Movement.   |



|    |  |   |
|----|--|---|
| 7  | Narrow search by dates                   | You are writing a paper about uranium mining on the Navajo reservation. You need to cite sources on this topic that were published recently. Use OneSearch to find articles and limit your search results to the last five years.   |
| 8  | Narrow search by discipline              | You are writing a paper about uranium mining on the Navajo reservation. Your assignment requires that one of the sources for your paper be from a historical perspective (i.e. written by a historian). Use OneSearch to find a book or article that you think you could use. |
| 9  | Find encyclopedia entries                | You are writing a paper about uranium mining on the Navajo reservation. You don't know much about this topic and want to get some background information about it. Use OneSearch to find an encyclopedia entry on the topic.  |
| 10 | Finding and requesting a book using CSU+ | You're writing a paper on mudejar architecture and you notice that there aren't many books at the CSUMB Library on your topic. You've heard that there's a way to get books from other CSU libraries. Find and request a book on your topic.                                  |
| 11 | Finding an ebook on a topic              | You're doing research on lone wolf terrorism, and you want to get started reading even though you're off campus for Spring Break. Find and start reading an ebook on your topic.  |
| 12 | Find an issue of a journal               | Your instructor has told you to read an article by author Wonhyung Lee in the journal <i>Nonprofit Management &amp; Leadership</i> . She can't remember the article title, but she said it's in the Fall 2017 issue. See if you can find the article.                         |

## Appendix 2: Results of Z-test for two proportions

### A/B test 1:

- Does the order of format types (alphabetical or frequency) have an effect on how often the “book” format is chosen?
- H0 (null hypothesis): the percent of clicks on the book format is the same regardless of the order of format types.
- H1 (alternate hypothesis): the percent of clicks on the book format varies depending on the order of format types.
- The Z-Score is -1.6637. The p-value is 0.09692. The result is not significant at  $p < 0.05$ . The null hypothesis cannot be rejected.
- In the format facet, listing format options alphabetically or in terms of frequency had no effect on how often the “book” format was chosen.

### A/B test 2:

- Does the location of the subject/topic facet have an effect on how often it is used when it is collapsed?
- H0 (null hypothesis): the percent of clicks on the subject facet is the same regardless of its placement.
- H1 (alternate hypothesis): the percent of clicks on the subject facet varies depending on its placement.
- The Z-Score is -0.4264. The p-value is 0.6672. The result is not significant at  $p < 0.05$ . The null hypothesis cannot be rejected.
- When the subject/topic facet is collapsed, having it at the top versus lower down in the facet list does not affect how often it is used.

### A/B test 3:

- Does the number of default results affect the number of clicks on “load more results”?
- H0 (null hypothesis): the percent of clicks on "load more results" is the same regardless of the number of default results.
- H1 (alternate hypothesis): the percent of clicks on "load more results" varies depending

on the number of default results.

- The Z-Score is -4.5256. The p-value is 0. The result is significant at  $p < 0.05$ . The null hypothesis is rejected.
- Users are more likely to click “load more results” when the initial results list includes 20 results than they are when it includes 10.

A/B test 4:

- Does having the subject/topic facet expanded to the first five results affect how often users click on it when this facet is lower in the facet list?
- H0 (null hypothesis): the percent of clicks on the subject facet is the same regardless of whether it is expanded.
- H1 (alternate hypothesis): the percent of clicks on the subject facet varies depending on whether it is expanded.
- The Z-Score is -4.3891. The p-value is 0. The result is significant at  $p < 0.05$ . The null hypothesis is rejected.
- When the subject facet is lower down on the facet list, having it expanded to the first five results results in greater use.

A/B test 5:

- Does the location of the subject/topic facet affect the frequency of use when this facet is expanded to view the first five results?
- H0 (null hypothesis): the percent of clicks on an expanded subject facet is the same regardless of its placement.
- H1 (alternate hypothesis): the percent of clicks on an expanded subject facet varies depending on its placement.
- The Z-Score is -7.199. The p-value is 0. The result is significant at  $p < 0.05$ . The null hypothesis is rejected.
- When the subject/topic facet is expanded (to view first five results), having it at the top of the facet list (versus lower down) results in greater use.

A/B test 6:

- Does the placement of the “send to” options affect their frequency of use?
- H0 (null hypothesis): the percent of clicks on the "send to" options is the same regardless of their placement on the page.
- H1 (alternate hypothesis): the percent of clicks on the "send to" options varies depending on their placement on the page.
- The Z-Score is 3.6786. The p-value is 0.00024. The result is significant at  $p < 0.05$ . The null hypothesis is rejected.
- Users are more likely to use the “send to” options when they are located at the top of the page versus the bottom.

A/B test 7:

- Does the placement of the “sign in” bar affect the proportion of users signing in?
- H0 (null hypothesis): the percent of sign ins is the same regardless of its placement on the page.
- H1 (alternate hypothesis): the percent of sign ins varies depending on its placement on the page.
- The Z-Score is 0.0914. The p-value is 0.92828. The result is not significant at  $p < 0.05$ . The null hypothesis cannot be rejected.
- The percent of users who sign in is the same regardless of whether the sign in bar is at the top of the page or the middle of the page.