

Comparison of University Course Search Interfaces

Andrew E. Munoz, Xiang Li, Bahadir A. Pehlivan, Frederick C. Harris, Jr., and
Sergiu Dascalu

Department of Computer Science and Engineering
University of Nevada, Reno
Reno, Nevada, U.S.A. 89557
{amunoz24, xli, bpehlivan}@nevada.unr.edu
{fred.harris, dascalus}@cse.unr.edu

Abstract

Universities must implement interactive and useful search tools that allow students to search and find various courses that they wish to enroll. In this study, a comparison is made between three different University course search tools. These include the University of Nevada, Reno, Harvard University, and University of California, Berkeley. Participants in the study executed various tasks on each website and then compared each website based on ease of use and efficiency. While each web page has different strengths, the results showed that many preferred certain design elements over others such as a more simplistic design with a clear layout. This paper details the various comparisons made between the websites through experimental tasks as well as how data was gathered, analyzed, and processed.

1 Introduction

This project is focused on the comparison of online course search interfaces of several different universities. University websites provide students with the ability to: search for a certain course, search courses by given specifications, and get information about a course such as required textbooks, syllabus or seating capacity. Comparisons of these different interfaces was made by measuring how easy, intuitive, and efficient it is for the user to navigate through the websites. In relation to human computer interaction, this project sought to understand different aspects of web search interface design through user performance and feedback.

University course search interfaces are very important because they provide the primary method for University students to find and register for their courses. This experiment benefits Universities and students because it allows for analysis of good design and what is more beneficial to the overall user experience. The main objectives of this project are to compare the usability and user interface of each website and use those results to determine what design elements leads to the optimal user experience. Overall, this project is designed to compare different course search design methods and determine through a user study the essential design techniques that relate to efficiency and ease of use.

The paper is organized as follows. In Section 2 similar work is briefly mentioned. Section 3 gives details about the experiments conducted, specifically about participants in 3.1, apparatus used in experiments in 3.2, how experiments are carried out in 3.3, which tasks were assigned to participants in 3.4 and general design of the experiments in 3.5. After this, Section 4 explains both the qualitative and quantitative results which are acquired from experiments and details Analysis of Variance (ANOVA) from the data. Lastly, conclusions are made and possible future work is discussed in Section 5.

2 Related Work

University course search web pages are very important because students and faculty will consistently interact with these web pages and therefore navigation needs to be efficient and intuitive. In this section, we go over some related works such as studies of specific websites or works that build new models, frameworks, or tools to analyze websites.

One such study goes over an enhancement of a technique called Fuzzy Analytic Hierarchy (FAHP) and the use of several different decision makers to obtain various results in evaluating the effectiveness of websites [1]. Another study presents a multi-criteria model, which is a hybrid model that uses multiple group decision making approaches, to use in the evaluation process of the quality of multiple university websites [2]. A third paper discusses the development of a comparison model that uses multi-criteria decision making algorithm AHP [3]. Each of these studies are related as they all utilized an advanced statistical model to analyze usability of websites.

There are also studies that focus on the examination of specific University websites and their usability. A paper by Ahmet Mentis [4] evaluates the Namik Kemal University (NKU) website and their findings provide some instruction to develop a more usable and efficient web page. Two other studies both compare three different University websites, similar to our study, with one focusing on specific usability problems [5] and the other focuses on a usability assessment that goes over task success rate, task completion times, and feedback [6].

3 Methodology

The following section is broken up into five subsections, each of which detail different aspects of the project. The subsections include Participants, Apparatus, Procedure, Tasks, and Design. In these subsections, the user study will be explained in detail as well as the tools that were used to accomplish the goals of the study.

3.1 Participants

For this study, there was a total of twelve participants in the age range of 20 to 30 years old. Half of the participants were students from the University of Nevada, Reno and the other half were non students. This allowed us to see how both experienced and inexperienced users adjusted to the course search websites and whether experience was necessary to navigate the websites efficiently. In order to get participants, we asked individuals if they would like to participate in the study until we reach a total of twelve participants. The participants were familiar with computers while also coming from different backgrounds. Compensation was not be provided to the participants for taking part in the study. The experiment took approximately 30 minutes

for each participant to complete. They were asked to complete a questionnaire at the beginning and end of the study.

3.2 Apparatus

The testing environment was in a quiet room with soft fluorescent lighting and white painted walls. The room had multiple computers and no visual distractions. The system that was used for this study was on a Dell XPS 8700 Desktop computer. This desktop has an Intel Core i7-4770 CPU with a XPS mini tower that stands at 16.01" high [7]. It also has 12GB of RAM along with a ZOTAC GeForce GTX 1080 GPU. Figure 1 shows a screenshot of the hardware used in this study, which includes two 23" monitors with mouse and keyboard.



Figure 1: Shows image of Dell XPS 8700 desktop with mouse and keyboard that were used to conduct experiment.

The software that was used to gather quantitative experimental data is called Mousotron. This software is a mouse and keyboard usage measurement tool. Mousotron keeps track of how far the mouse has traveled on the screen, number of keystrokes, cursor speed, mousewheel scrolls, and number of clicks [8]. Figure 2 displays the Mousotron interface. It was used in this project to keep track of left and right mouse clicks as well as double clicks the user had while navigating each website. It was also utilized to measure the number of keystrokes and time spent accomplishing the task.

3.3 Procedure

Experiment starts with an entry questionnaire which asked participants:

- Their Age and Gender
- Whether they are currently enrolled in any University/College
- Level of familiarity with each of the websites

Following the questionnaire, the participants were each allowed a two minute demo for each website in order to quickly familiarize themselves with each of the University websites. During



Figure 2: Display of the vertical Mousotron interface.

demos no specific task was assigned to participant, instead they were simply told to get familiar with the website. No training or tutorial were given. Upon completion of the website demos, participants started to do their first task for the assigned website. Due to the simple nature of the tasks, participants were only verbally instructed about what to do for each task. Generally, participants were able to complete tasks in a relatively short time frame, therefore there were no breaks with the exception of a supervisor resetting the logger software between tasks. After all tasks of a website were finished, the experiment continued with the first task of the next assigned website. After participants finished all the tasks, they were given an exit questionnaire which asked participants:

- To rate ease of course search and navigation for each website
- To rate ease of finding information about courses for each website
- To rank all websites and provide general satisfaction with each design
- Their comfort level with the apparatus

3.4 Tasks

Generalizability was considered when designing tasks for this study. Generalizability is important for external validity [9]. These tasks were chosen to mimic the process a typical student goes through when signing up for classes. The first task asked the participant to look up a course by the name of the course and record the professor's name. Course names varied for each university. Course look up order was chosen randomly between the three websites and were for the Spring 2019 semester across all universities. The second task asked the user to look up a course by the course code and record the professor's name. The first two tasks were chosen to compare two common methods students can use to find courses. The third task asked the user to look up two specified courses and determine if there were conflicting lecture meeting times. Tasks were evaluated using the metrics defined in Section 3.5.

3.5 Design

The independent variable in our study was university course search websites. Websites from three different universities were evaluated. These websites were from: The University of Nevada,

Reno, Harvard University, and The University of California, Berkeley. Figures 3, 4, and 5 display screenshots of search results pages from each of the three University course search websites. Each participant was asked to perform the tasks described in section 3.4 for each website. Counterbalancing was required in order to minimize variance originating from the order of websites evaluated [9]. Without counterbalancing, participants could have better performance with latter websites evaluated as they get used to navigating these websites and become familiar with the tasks asked of them. Alternatively, participant performance could decrease due to fatigue for latter websites. With three factor independent variable, there were six possible permutations of order. Two participants were assigned to each permutation for a total of 12 participants.

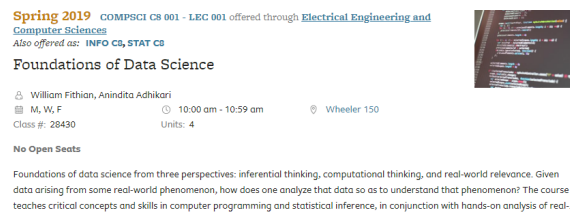


Figure 3: Screenshot of results page from UC Berkeley.

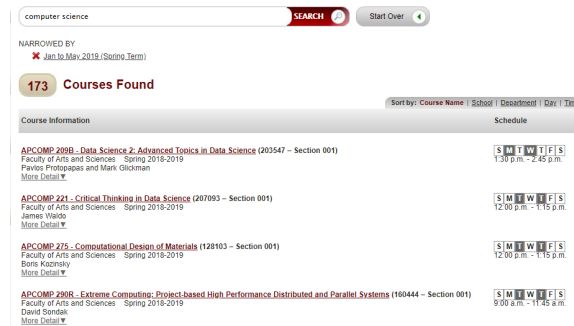


Figure 4: Shows results from a search on Harvard University.

The dependent variables were grouped into two classes: user satisfaction and performance metrics. User satisfaction is a subjective measure while performance provides objective measures to compare differences in our independent variable [10]. User satisfaction was measured using a survey that asked several questions related to ease of navigation and user satisfaction of design. Following the work of Hasan [11], study participants were asked to rate satisfaction on a numbered scale. This method allowed semi-quantitative analysis of satisfaction. Performance was measured by comparing task completion time, number of mouse clicks, distance cursor traveled, and mouse scroll wheel movements. We hypothesize that number of left mouse clicks would be a good proxy for ease of use. A more intuitive and efficient web page should require less left mouse clicks to navigate to the information requested. The less amount of mouse wheel movement should also indicate how easy it is to find the information. A web page with an overwhelming amount of information per page would require the user to use more mouse wheel movement. In addition, more scroll wheel movements could indicate poor specificity of the

search engine, thus leading to more search results. A well designed web page should have intuitive link and button placement, this is measured through cursor movement distance. Better web pages should have less cursor movement. Task completion time should be lower for easier to use web pages. Analyzing multiple types of measurements of the same activity is called triangulation [12]. This method strengthens and serves to validate the data collected.

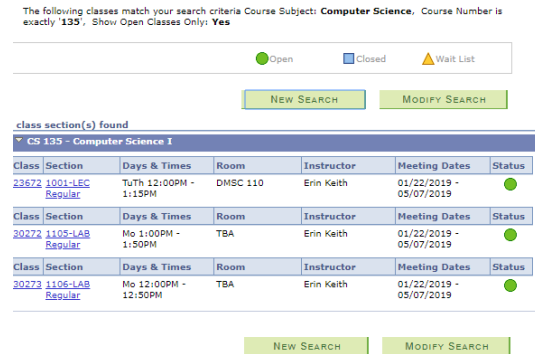


Figure 5: Display of University of Nevada, Reno’s (UNR) course search results page.

Our main control variable were the tasks that participants completed which were the exact same across all participants. The second main control variable was the testing apparatus; this was consistent for each participant. However, a random variable that is associated with the testing apparatus is familiarity with the testing instrument. Increasing familiarity might lead to better performance, however, all participants reported high familiarity with the testing station in the post experiment questionnaire. Mental fatigue is another random variables we have to consider. Performance would be less efficient as the user’s cognition is affected by fatigue. Other random variables stem from the environment so we tried our best to control these by offering a quiet testing environment that is free from outside distractions. Variance from user to user differences would introduce unnecessary noise to our quantitative analysis of performance.

4 Results and Discussion

4.1 Mouse Left Clicks

Number of mouse clicks showed a big difference for UNR’s website. Its effect can be seen in all three tasks. Figure 6 shows box plots of mouse left clicks for each of the three tasks. Even though standard deviation is large for UNR, it is still obvious from Figure 6 that UNR required far more left clicks to complete any task. The large variation was probably caused by the wide range of prior experience with UNR’s web page. Half of the participants were UNR students who were were already familiar with that web page. The other half had never attended any of the three universities, therefore their mouse left click count were more indicative of intuitiveness. Even though this large variance exists, analysis of variance (ANOVA) for the first task showed significance for left mouse clicks ($p = 0.0007$) as well as for the second task ($p = 0.0054$) and the third task ($p = 0.0007$). A post hoc Scheffé test revealed statistical significant differences between the UNR and Berkeley pairing as well as the UNR and Harvard pairing for all three tasks. There were no significant differences found between Harvard and Berkeley.

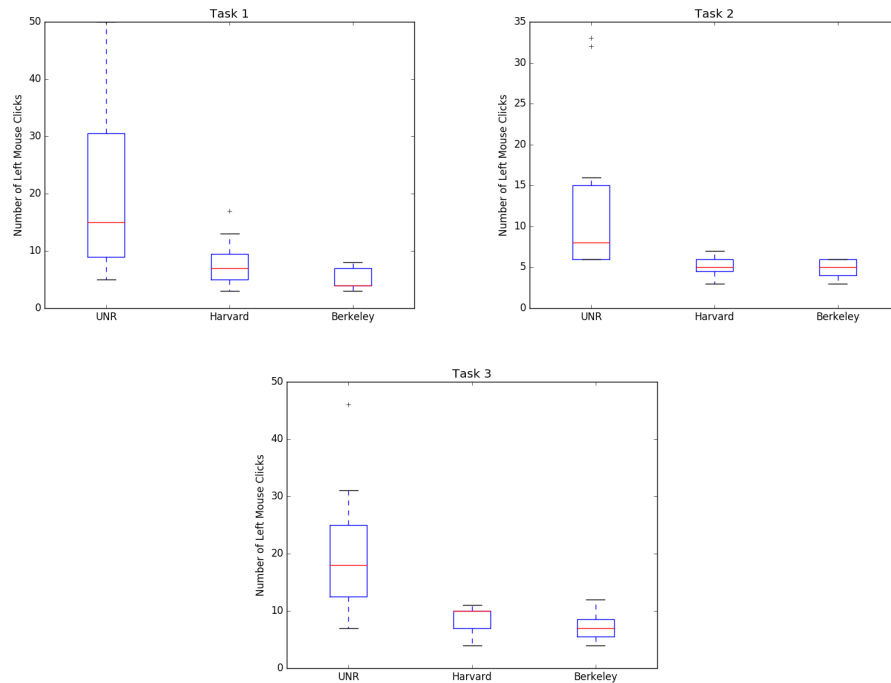


Figure 6: Graphs that display average number of left mouse clicks for the three tasks.

We observed that participants usually had a hard time navigating UNR's website to complete the tasks. There was an overwhelming amount of text data entry boxes and other options selections compared to the simple design of the other two universities. As seen in Figure 6, there was not much difference in left mouse clicks between task one and task two for Berkeley and Harvard, but there was a large difference for UNR's task one and task two. The UNR web page did not implement course search by name as well as the other two universities. Even though the UNR page was optimized for search by course code, it still required more left mouse clicks for the second task. Also, the UNR web page required participants to input search terms in a very specific format and perfect spelling, often times returning no search result at all. The search engines of the other two websites were more robust.

4.2 Mouse wheel Movement

The number of mouse wheel movements per task was high for the Berkeley course website as shown in Figure 7. This was due to participants needing to scroll more to find the requested courses in Berkeley compared to other websites. The Berkeley search results page displayed more information per course and in bigger format compared to the two other universities. ANOVA revealed no significant differences for task one and three. Task two was significant with ($p = 0.0176$). Post hoc tests revealed significant differences between UNR-Berkeley and Berkeley-Harvard. Harvard's web search returned more specific results page with fewer results in a compact format, so it required no mouse wheel scrolling.

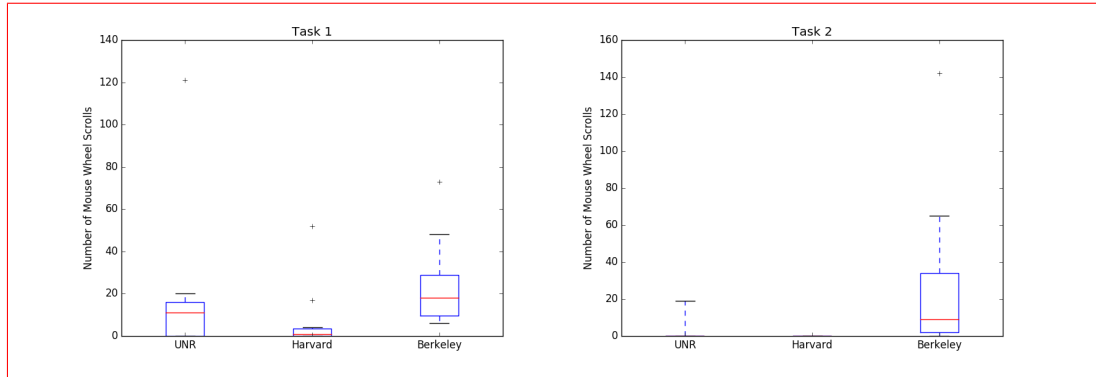


Figure 7: Average mouse wheel movements for all participants. Harvard resulted in no use of the mouse wheel for task 2 and only one participant used the mouse wheel when navigating UNR.

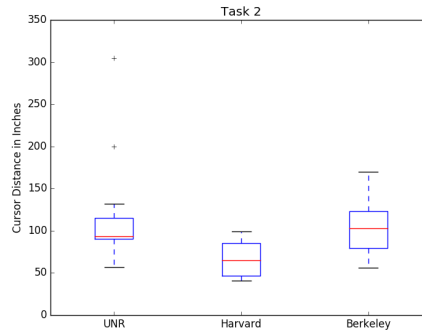


Figure 8: Displays the average cursor distance travelled for all participants.

4.3 Cursor Movement Distance

Statistical significance was only found for cursor movement distance for task 2 ($p = 0.0141$). Scheffé post hoc test revealed significant difference in the UNR-Harvard pairing. Harvard's website had less cursor movement distance than the other two universities as seen in Figure 8. This suggests that the Harvard website had a more efficient design, requiring the user to move the cursor less. Variances for these measurements were relatively high, which is why there wasn't any significant differences found by ANOVA for other tasks. Some of this variance stems from participant cursor movement behavior. Certain users naturally move the cursor more or less than others. In general, UNR had higher variance across all three tasks. This extra variance is most likely due to difficulties in locating correct entry boxes and options on the UNR page for participants who had no prior experience.

4.4 Task Completion Time

Task completion time had significant differences for task one ($p = 0.003$) and task three ($p = 0.03$) only. Post hoc Scheffé revealed significant differences between UNR-Harvard and UNR-

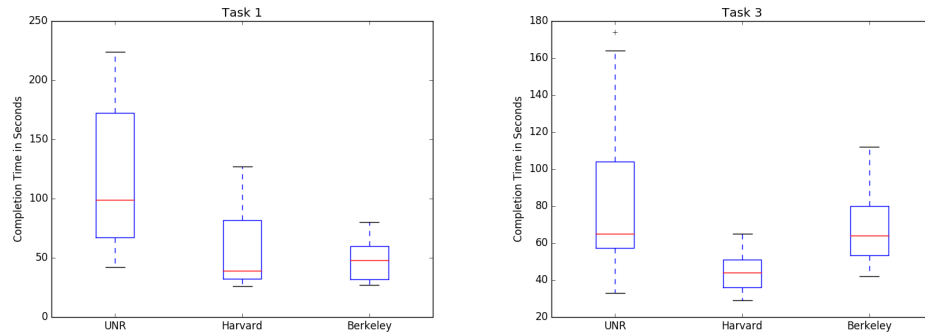


Figure 9: Shows average task completion times for Task 1 and 3.

Berkeley for the first task, but only UNR-Harvard for the third task. UNR required more time to complete tasks as shown in Figure 9. The search bar employed by both Harvard and Berkeley was successful in reducing completion time. The UNR page had a confusing search system that required more navigation to select the correct subject, or prior knowledge of the course code. Also the keyword search option was hidden and didn't work well as was described in Section 4.1. The UNR web page took a much longer time to load search results. This was indicated by a pinwheel on the left side of the screen, which most participants did not notice causing them to click the search button multiple times. In terms of number of key presses no significant results were obtained to compare websites. Graphs shown in Figure 9 show completion time for task one and three.

4.5 Qualitative Results

In the post questionnaire, eight participants ranked Berkeley as having the best experience, while four ranked Harvard as having the best. No respondent preferred the UNR page over the other two. In terms of user satisfaction and ease of navigation/use, most respondents gave Berkeley and Harvard web pages higher ratings (4-5 on our likert scale) compared to UNR. There were three participants who gave UNR above a three on ease of navigation, but only one person gave above a three on satisfaction. Most participants preferred the design of the Berkeley and Harvard pages, regardless of whether they were UNR students or not. They felt that the Harvard and Berkeley web pages provided a more satisfying user experience and they were easier to find information and navigate. All participants responded as comfortable with the testing setup.

5 Conclusion And Future Work

We carried out experiments to compare three different University websites in terms of their usability and design. In comparing the websites of Harvard, UNR, and UC Berkeley we found that users tended to prefer the design of the search tool for both Harvard and UC Berkeley while they were mostly dissatisfied with the UNR search tool. We obtained these results by considering the number of left mouse clicks per task, cursor movement, mouse wheel scroll, and completion time. These metrics revealed that the Harvard and Berkeley design enabled better

user performance. User feedback was also gathered through a questionnaire to better understand the user thought process while using the websites and also to gain an insight into good design techniques. These questionnaires revealed that all participants preferred the Harvard and Berkeley course web pages to the UNR page.

Further work involves expanding this study to include more participants, especially those with no experience with the UNR website, so that our results will have more statistical power and less bias. The dependent variables that exhibited no significant differences between the websites could gain meaning with more participants, enabling the extraction of more information. From this study, we can design guidelines that instruct how to make course websites more user friendly and efficient. In order to accomplish this, we can examine specific website components (e.g. the search bar) to see what design choices users find more intuitive as well as visual aesthetics. Using this knowledge, it would be useful to implement these design changes and perform a subsequent user study to investigate their effects.

References

- [1] R. Nagpal, D. Mehrotra, P. K. Bhatia, and A. Sharma, "Rank university websites using fuzzy ahp and fuzzy topsis approach on usability," *International journal of information engineering and electronic business*, vol. 7, no. 1, p. 29, 2015.
- [2] D. Pamučar, Ž. Stević, and E. K. Zavadskas, "Integration of interval rough ahp and interval rough mabac methods for evaluating university web pages," *Applied Soft Computing*, vol. 67, pp. 141–163, 2018.
- [3] S. Roy, P. K. Pattnaik, and R. Mall, "Quality assurance of academic websites using usability testing: an experimental study with ahp," *International Journal of System Assurance Engineering and Management*, vol. 8, no. 1, pp. 1–11, 2017.
- [4] S. A. Mentis and A. H. Turan, "Assessing the usability of university websites: An empirical study on namik kemal university." *Turkish Online Journal of Educational Technology-TOJET*, vol. 11, no. 3, pp. 61–69, 2012.
- [5] L. Hasan, "Heuristic evaluation of three jordanian university websites." *Informatics in Education*, vol. 12, no. 2, pp. 231–251, 2013.
- [6] S. Roy, P. K. Pattnaik, and R. Mall, "A quantitative approach to evaluate usability of academic websites based on human perception," *Egyptian Informatics Journal*, vol. 15, no. 3, pp. 159–167, 2014.
- [7] Dell Inc, "XPS 8700," first accessed December 12, 2018. [Online]. Available: <https://www.dell.com/en-us/work/shop/cty/pdp/spd/xps-8700>
- [8] Blacksun Software, "Mousotron: Mouse and Keyboard Usage Measurement," first accessed November 3, 2018. [Online]. Available: <http://www.blacksunsoftware.com/mousotron.html>
- [9] I. S. MacKenzie, *Human-Computer Interaction An Empirical Research Perspective*, 1st ed. Waltham, MA: Elsevier, 2012.
- [10] T. Jokela, J. Koivumaa, J. Pirkola, P. Salminen, and N. Kantola, "Methods for quantitative usability requirements: a case study on the development of the user interface of a mobile phone," *Personal and Ubiquitous computing*, vol. 10, no. 6, pp. 345–355, 2006.
- [11] Layla Hasan, "The Usefulness of User Testing Methods in Identifying Problems on University Websites," *Journal of Information Systems and Technology Management*, vol. 11, no. 2, pp. 229–256, aug 2014.
- [12] J. Emanuel, "Usability testing in libraries: methods, limitations, and implications," *OCLC Systems & Services: International digital library perspectives*, vol. 29, no. 4, pp. 204–217, oct 2013. [Online]. Available: <http://www.emeraldinsight.com/doi/10.1108/OCLC-02-2013-0009>