Heuristic Evaluation and Usability Testing

Saki Takizawa
Earlham College
Richmond, Indiana
stakiz21@earlham.edu

ACM Reference Format:
Saki Takizawa. 2025. Heuristic Evaluation and Usability Testing . In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 ABSTRACT

User experience (UX) design is a critical approach to web design that emphasizes both the aesthetic and functional aspects of a website. However, not all websites possess well-designed UX. Heuristic evaluation and usability testing serve as two valuable methods to enhance it. The evaluations in this paper aim to pinpoint usability issues on a college website and formulate guidelines for its redesign.

This paper will employ two research methods, heuristic evaluation and user testing, to identify and sort out problems affecting the user experience of a college website. The evaluation will create a list of potential usability issues by referencing Nielsen's heuristic evaluation principles and the Research-Based Web Design and Usability Guidelines provided by the US Department of Health and Human Services [6][8]. Subsequently, the college website will undergo heuristic evaluation based on this list. Following this step, tasks for the user testing phase will be derived from the identified issues or potential concerns. Ultimately, the results and analysis from the user testing will shed light on specific areas of the website in need of improvement. The overarching goal is to enhance the website's usability, making it more user-friendly and enjoyable for visitors.

2 INTRODUCTION

User Experience (UX) has become a central paradigm within the field of Human-Computer Interaction (HCI). The demand for well-designed digital interfaces has been steadily increasing, necessitating designs that are not just highly functional but also intuitive, enjoyable, and effective in meeting the diverse needs of users. Within this context, the evaluation of user experience is of paramount importance as it involves understanding the efficiency and user-friendliness of digital interfaces. One of the valuable methodologies for assessing the usability of a website is heuristic evaluation. According to Nielsen, "Heuristic evaluation is a usability inspection method for computer software that helps to identify usability problems in the user interface (UI) design. It involves having a small set of evaluators examine the interface and judge its compliance with recognized usability principles (heuristics) [8]."

In addition to heuristic evaluation, user testing, which engages actual users in interacting with a system, provides a holistic view of usability. This research explores the significant intersection of UX, heuristic evaluation, and user testing, with a specific focus on the usability of college websites. College websites serve as vital information resources for college students, faculty, prospective students, and even more.

In this research, I will begin by applying heuristic evaluation to the Earlham Computer Science Department's web page, referencing Nielsen's heuristic evaluation principles and the Research-Based Web Design and Usability Guidelines provided by the US Department of Health and Human Services [6][8]. Subsequently, I will catalog the problems identified in the website and rank them based on the severity of the design issues [3].

The next step involves creating usability scenarios that will be implemented using a remote usability testing tool. In this research, Loop11 will address the usability testing. Following this, user testing will be conducted with different user groups employing the remote usability testing tools. Finally, after collecting the data from the user tests, we will determine which parts of the website require improvement by incorporating Analysis of Variance (ANOVA).

This research stands out for its unique combination of several different evaluation methods, a focused analysis specialized to the domain of college websites, and a data-driven redesign strategy. All of them aimed at significantly enhancing the user experience for college website visitors.

3 BACKGROUND

Heuristic Evaluation and user testing play pivotal roles in UX design. UX design is an approach that aims to produce a better user experience. It encompasses the entire process of creating a product, including aspects of branding, design, usability, and function [1]. UX design involves two elements: how a person perceives and how a product is used. Both of these elements depend on each other in UX design.

Perception in UX design encompasses visual aspects, such as layout and the use of colors, emotional responses, such as how users rate their overall experience with the website, and comprehension, such as how easily users can understand the website's layout. Additionally, usage in UX design, focusing on how a product is used, includes navigation, such as how easily users can move from one section to another on the website, functionality, such as the performance of functions like buttons, and task efficiency, assessing how effectively users can find the information they were seeking.

These elements are intricately linked in UX design, highlighting the significance of designing products with UX principles that

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

effectively capture and respond to human emotions and experiences. Adhering to these principles is essential for creating products that meet user needs and preferences.

3.1 Heuristic Evaluation

Heuristic evaluation is a usability evaluation method that involves a systematic inspection of a user interface by usability experts against a set of predefined usability principles. The evaluators identify usability problems, such as violations of these principles, and provide recommendations for improving the interface's overall usability. The aim of heuristic evaluation is to uncover usability issues quickly and cost-effectively, ensuring that the interface is more user-friendly and efficient in meeting the needs of its intended users. Thus, heuristic evaluation is a way to test a user interface for usability problems by comparing it to a set of best practices.

Heuristic evaluation is a useful method for identifying usability problems, but it is important to be aware of its limitations [5]. Experts may identify problems that users do not encounter, or they may miss problems that users are having. This is because heuristic evaluation is based on experts' knowledge and experience, which may not always be aligned with the needs of the target users. Additionally, heuristic evaluations are typically conducted in a lab setting, which may not be representative of how users interact with the product in the real world. Finally, heuristic evaluations can be subjective, and different experts may identify different problems. To mitigate these limitations, it is important to use heuristic evaluation in conjunction with other usability testing methods, such as user testing.

3.2 Usability Testing

Usability testing, particularly in the realm of UX (User Experience), is a method employed to assess the ease of use of a web application. This assessment is carried out by instructing users to perform specific tasks with the application. Usability testing can take various forms, including in-person sessions where a researcher observes users completing tasks, and remote testing, which offers a more accurate perspective on real-world user experiences [4].

Compared to heuristic evaluation, usability testing tends to unearth more significant issues. One distinct advantage of usability testing is its ability to pinpoint problems that could affect actual users of the application, without the need for pre-sorting or filtering these issues based on their perceived impact. The test itself helps in assessing the impact of identified problems. This has been consistently demonstrated in Jeffries and Desurvire's studies, where nearly all problems identified through usability testing proved to be of above-average severity. In addition, the problems that usability tests unearth are typically not found in other methods, such as the heuristic evaluation, because the variation of actual user's actions from users surpass the level of expectation of experts in most of the cases [7].

In the case study at Boğaziçi University, one of the user groups that the researchers carried out the users test on the undergraduates at Boğaziçi University. The ten tasks were like following: ~~Sure,~~ here are the paraphrased versions of each task from the case study at Boğaziçi University:

(1) Explore the website of the laboratory related to Flexible Automation. Guideline 10.4: Avoid misleading cues to click [6].
(2) Locate the contact number for the chairman of the Industrial Engineering(IE) department. Guideline 2.5: Design for working memory limitations [6].
(3) Find the webpage for the Office of International Relations to gather information about Erasmus or Exchange programs. Guideline 10.4: Avoid misleading cues to click [6].
(4) Search for the courses taught by IE Professor Prof. Dr. Barbarosoğlu. Guideline 10.4: Avoid misleading cues to click [6].
(5) Navigate to the webpage of the "Quantitative Finance Research Group" within the Boğaziçi University Industrial Engineering Dept. Guideline 10.4: Avoid misleading cues to click [6].
(6) Locate the information page to determine if PSY 101 is available as an HSS elective. Guideline 16.2: Structure each content page to facilitate scanning [6].
(7) Find the Alumni list of the IE Department to connect with fellow graduates after completing your studies. Guideline 10.4: Avoid misleading cues to click [6].
(8) Find the contact number for Instructor Dr. Yasemin Aksoy to inquire about a specific elective course. Guideline 2.5: Design for working memory limitations [6].
(9) Check if Dr. Suat Genç offers any undergraduate courses as elective options due to his reputed success as an instructor. Guideline 16.4: Group related elements [6].
(10) Discover the list and descriptions of IE-Elective courses available for undergraduate education. Guideline 16.2: Structure each content page to facilitate scanning [6].

Addition to the tasks above, users would be asked about demographics and the frequency of internet usage to see if there is any significant difference in those categories as the case study at Boğaziçi University did.

3.3 Analysis of Variance

After collecting the data from usability testing, Analysis of Variance (ANOVA) will be introduced to investigate the results. ANOVA is used to determine if there is a statistically significant difference between user groups.

In simpler terms, ANOVA is a statistical test that can be used to compare the means of three or more groups. It is a powerful tool, but it is important to note that it has three key assumptions: the samples must be independent, the variances of the groups must be equal, and the data must be normally distributed.

In the case study, Mahmut Ekşioğlu and colleagues conducted additional comparison tests alongside ANOVA to ensure that their results were reliable [3]. This is a good practice to follow, as it helps to minimize the risk of making false conclusions.

This research will use ANOVA, a statistical tool, to dig deeper into how different groups of users interact with the college website. It aims to compare how well different groups perform specific tasks and assess their overall experience. For example, we'll see if there are significant differences in how well different groups, like prospective students, faculty, and current students, complete tasks.

Finding these differences can help us identify areas of the website that need to be improved for specific user groups. For instance, if prospective students have much lower success rates in navigating or finding information, it suggests that the website might not be meeting their needs. By using ANOVA to pinpoint these differences, we can make targeted improvements to better address the needs of specific user groups and improve the overall user experience of the college website.

4 DESIGN AND IMPLEMENTATION

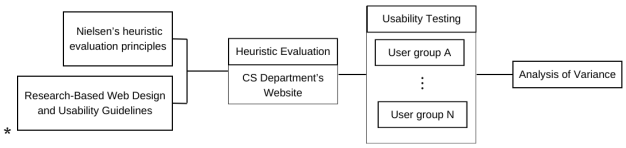

Figure 1: My Reserach Plan

This research will involve four main steps. First, a guideline will be created to evaluate the Earlham College Computer Science Department website. This guideline will be based on Nielsen's heuristic evaluation principles [9] and Research-Based Web Design and Usability Guidelines. To ensure that the website's integrity is not affected by any updates during the research, the current version of the website will be cloned. These references will identify potential usability issues from a human perspective. Using these guidelines, areas of the website that require improvement or potential redesigns in terms of usability will be identified.

4.1 Heuristic Evaluation Principles

This research carefully selects key principles and guidelines to evaluate the website's usability. These come from two well-respected sources: Nielsen's Heuristic Evaluation Principles and the Research-Based Web Design and Usability Guidelines. To pick the most important ones, we looked for elements marked with stars in these resources. These starred points are considered especially important because they are based on fundamental principles of user-centered design and have a big impact on how people experience the website. By focusing on these crucial points, our evaluation will primarily identify the areas with the biggest potential for improvement. The evaluation guideline will look like the following.

(1) Understand and Establish User's Expectation
   - If the website provides contents that are appropriate.
   - Make sure that the website includes what is user's requirement
(2) Avoid having unwanted windows and graphs
   - Unsolicited windows and graphs are not necessary.
(3) Avoid using color alone to convey call-to-action
   - Consider that all people can recognize colors in the same way so only changing color to convey the critical actions is not effective.
(4) Enable to have easy access to the homepage
   - Make sure that user can access to the homepage from any page in the website
(5) Display all the major options on the homepage
   - Show the major options that users take should be easily accessible
(6) Create positive impression at a first sight/ homepage
   - The homepage is a key for users to know the quality of the website so it needs to make a good impression.
(7) Cluttered layout is unwanted
   - Create pages that are not considered cluttered by users, such as cramped design.
(8) Place critical items consistently
   - The Important Information needs to be display all the time at the top center
(9) Avoid horizontal scroll
   - Designing appropriate not to have a horizontal scroll.
(10) Clear category labels
   - Make sure that all the category labels and their links reflect the information and items in the category.
(11) Buttons need to be distinct
   - Ensure that push buttons are indicated distinctively.
(12) Clear category labels
   - Make sure that all the category labels and their links reflect the information and items in the category.
(13) Text color and Background
   - Text color needs to be distinct from the background color. The ideal color is to use black text on the white background.
(14) Speak the users' language
   - The wording on the web page needs to be expressed clearly
   - Words, phrases, concepts on the website should be familiar to users
(15) Minimize the user's memory load
   - The information that people do not need to memorize in longer term, such as instructions to show how to use the page, should be simple.
   - Those information needs to be retriable easily.
(16) Consistency
   - Users never need to wonder if the different actions, situations, words mean the same.
   - Consistency also needs to be kept between the main page and sub pages.
(17) Provide clear exit pathways
   - All the situations that are possible on the web page have to have a clear exit pathway, so that users can leave the unwanted state while using the web page.
(18) Provide shortcuts
   - To avoid making users feel troublesome because it takes so many steps to get information, there needs to be some shortcuts to lead users to get what they need relatively easy and fast.

Using the list above, an evaluation will be conducted on the Earlham College CS Department Website [2]. For instance, the homepage contains a section where the text color and background color is not sufficiently contrasted. An item from the list, Text color and Background [6], can be applied to this case.

- Text color and Background
  - Text color needs to be distinct from the background color. The ideal choice is to use black text on a white background.

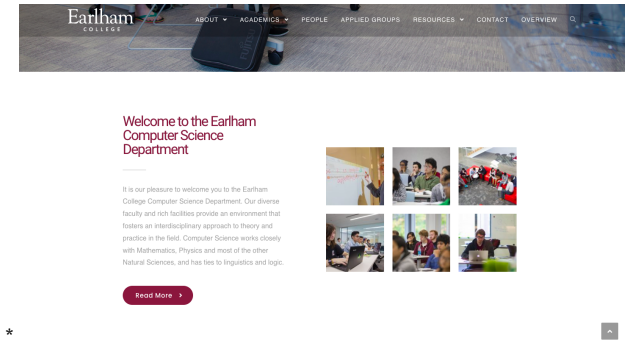

Figure 2: Original Screen

The presence of light gray text on a white background hampers readability for users, particularly for those with color blindness. As recommended in the item, employing black text on a white background is optimal. Since the website's background is already set to white and the title appears in maroon and a larger font size, using black for the body text would be appropriate. The change can be implemented as follows:

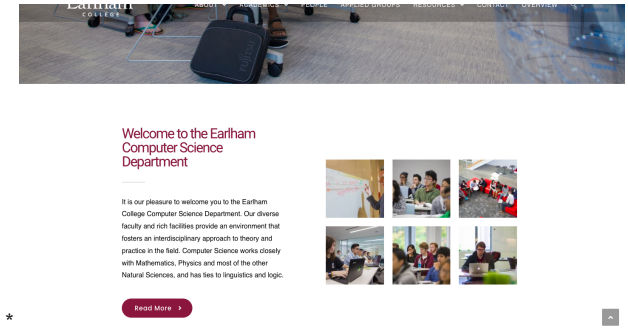

Figure 3: Revised Screen

Next, tasks will be created for different user groups to test the usability of the CS department website. These user groups will include current Earlham undergraduate students, prospective high school students, and faculty members of the CS department. The usability test will be conducted remotely in a survey-like format. The goal is to collect data on task completion, time taken for completion, as well as demographic information such as gender and ethnicity.

One of the risks in the usability testing phase is that maybe not enough data will be collected within the set time frame for the research, especially concerning high school students. Another risk is the wording in the usability testing. To avoid users confusion, it will be better to conduct a pre-survey with similar questions and smaller groups from the user pool to identify any potential ambiguities or challenges in the wording of the tasks.

The collected data will then be analyzed using the analysis of variances method. This analysis will identify which user group and tasks had the most significant impact on usability. Finally, a guideline for redesigning the website will be created based on the findings.

5 TIMELINE

(1) First month
- Clone the website
- Create a guideline for heuristic evaluation
- Evaluate and figure out where usability issues might happen
- Create a pre-survey draft based on the result from the heuristic evaluation

(2) Second Month
- Pre-survey with smaller number of users to finalize the wording in usability testing
- Collect data by sending my survey to the potential users

(3) Third month
- List the parts where usability issues were found during the heuristic evaluation and user testing
- Analyzing with ANOVA method to find the significant impact on users that affects their experience through the use of the website

(4) Fourth month
- Create the guideline for redesigning the web page
- Finalizing research paper

REFERENCES

[1] 2023. User Experience (UX) Design Your constantly-updated definition of User Experience (UX) Design and collection of topical content and literature. https://www.interaction-design.org/literature/topics/ux-design
[2] Earlham CS Department. [n. d.]. Home - Earlham CS. https://cs.earlham.edu/
[3] Mahmut Ekşioğlu, Esin Kiris, Burak Çapar, Murat N. Selçuk, and Selen Ouzeir. 2011. Heuristic Evaluation and Usability Testing: Case Study. *Department of Industrial Engineering, Boğaziçi University, Istanbul-Turkey* (2011).
[4] Interaction Design Foundation. 2016. What is Usability Testing? *Interaction Design Foundation* (2016).
[5] Bendoly Laura Goldman, Kate Haley. 2003. Investigating Heuristic Evaluation: A Case Study. *ERIC* (2003).
[6] Health and Human Services Department. 2006. Research-Based Web Design Usability Guidelines. *Health and Human Services Department* (2006).
[7] Robin Jeffries and Heather Desurvire. 1992. Usability Testing vs. Heuristic Evaluation: Was There a Contest? *SIGCHI Bull.* 24, 4 (oct 1992), 39–41, numpages =. https://doi.org/10.1145/142167.142179
[8] Jakob Nielsen. 1995. How to conduct a heuristic evaluation. *Nielsen Norman Group* 1, 1 (1995), 8.
[9] Molich Rolf and Nielsen Jakob. 1990. Improving a Human-Computer Dialogue. *Commun. ACM* 33, 3 (mar 1990), 338–348. https://doi.org/10.1145/77481.77486