International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

Exploring Common Usability Methods and Their Application in Website Optimization

Ahmed Salem Ahmed Nasser

Ph.D. Researcher, Department of Management Science, Dr. Babasaheb Ambedkar Marathwada University, India. Under the guidance of Dr. Arshad Moizuddin Razvi
Email: *ahmed4edu.s[at]gmail.com*

Abstract: ***Website usability is a prominent factor in determining user experience, interaction, and business success. With web platforms becoming more complex, the necessity for systematic usability testing becomes more critical. This paper discusses five commonly used usability methods - Heuristic Evaluation, System Usability Scale (SUS), Eye Tracking, A/B Testing, and Usability Testing - outlining their processes, strengths, weaknesses, and real-world applications. Based on recent research and best practices in the field, the paper identifies how every approach contributes in a distinct way to the detection of usability problems and the development of optimization strategies. The results point out that choosing the right evaluation approach is a matter of project objectives, budget, and the phase of development, and that strategic integration of multiple approaches can result in more user-friendly, user-centered web design.***

Keywords: Usability Testing, Website Usability, Web Performance, Websites

1. Introduction

Usability plays a major role in the success of websites and user experience, and UX is now a make-or-break factor for a website's success. As websites grow in sophistication and functionality, making them easy to use, how easily and effectively users can use a system has never been more important. Usability has a direct impact on user satisfaction, engagement, and conversion rates; therefore, it is an important area of focus for designers, developers, and businesses likewise. In order to assess and improve website usability, various methods have been developed and improved over the years. Each of these methods provides individual perspectives on user behavior with digital interfaces and assists in determining areas of improvement.

This paper discusses five common usability evaluation methods: Heuristic Evaluation, System Usability Scale (SUS), Eye Tracking, A/B testing, and Usability Testing. Referring to current literature and relevant books, each approach will be described and analyzed in terms of process, strengths, weaknesses, and real-world implementation in website optimization. Rather than presenting a separate literature review, this paper integrates relevant research within each method section to better highlight the practical use and impact of these techniques in real-world contexts.

Knowing when and how to use these usability methods is critical to designing easy-to-use websites that are intuitive, effective, and engaging. For example, SUS provides a standardized and quick way of measuring overall usability, whereas eye tracking offers deep insight into visual focus and user behavior. Expert-based reviews can be done through heuristic evaluation, while usability testing uses real users to complete tasks and evaluate their performance. A/B testing, in contrast, is concerned with data-driven comparisons across design alternatives.

Through a discussion of these usability methods, this paper seeks to present a detailed overview of their function in maximizing websites and assisting practitioners in selecting the appropriate method at the appropriate time to improve user experience.

2. Heuristic Evaluation

Heuristic Evaluation is a usability inspection method introduced by Jakob Nielsen and Rolf Molich in 1990. It involves assessing user interfaces - such as websites - against a set of established usability principles known as heuristics to identify potential usability issues. A small group of expert evaluators independently examine the interface as each is likely to identify different problems. Nielsen recommends using three to five evaluators to balance the benefits of diverse insights with the practicalities of the evaluation process. Heuristics are not strict rules that interfaces must comply with but rather general guidelines that help designers in adhering to fundamental usability principles. This method is applied in early stage of the designing phase. During the evaluation, individual evaluators document the issues they encounter, assigning severity ratings to each problem, these findings are then aggregated, matched to the violated heuristics, and compiled into a report that includes actionable recommendations for improvement [1].

Commonly, a Heuristic Evaluation is conducted based on Jakob Nielsen's 10 usability heuristics [2], which are listed below:

1) **Visibility of system status** – The system should always provide timely and appropriate feedback to users.
2) **Match between system and the real world** – Use concepts, language, and metaphors that are familiar to the user.
3) **User control and freedom** – Support undo, redo, and easy exits from unintended actions, allowing users to easily correct mistakes.
4) **Consistency and standards** – Interface elements should follow platform conventions and maintain internal consistency.

5) **Error prevention** – Design systems to minimize the occurrence of errors in the first place.
6) **Recognition rather than recall** – Reduce users' memory load by making elements, actions, and options visible.
7) **Flexibility and efficiency of use** – Support both novice and experienced users by offering shortcuts and customizable features.
8) **Aesthetic and minimalist design** – Avoid unnecessary content and interface elements that do not serve a purpose.
9) **Help users recognize, diagnose, and recover from errors** – Provide clear, concise error messages and suggest solutions.
10) **Help and documentation** – Offer accessible, task-oriented help and documentation when needed.

A study by Habib et al applied Heuristic Evaluation in combination with user feedback to assess the usability of the *KUKERTA Portal*, a website used by students at Universitas Riau for managing community service programs. The evaluation was conducted using Nielsen's 10 usability heuristics, where a group of expert evaluators systematically reviewed the interface to identify usability violations. In parallel, a survey of 98 student users was conducted to gather insights into their real-world experiences using the portal. The findings revealed critical usability issues, especially concerning flexibility and efficiency of use, user control and freedom, and error prevention. For instance, the portal lacked intuitive navigation, failed to support undo/redo actions, and had inadequate error validation. The evaluators assigned severity ratings to each issue, and the combined results from the heuristic and user feedback analysis were used to generate a comprehensive usability report. This included specific recommendations to address the identified problems, demonstrating the effectiveness of Heuristic Evaluation as a practical tool for website optimization when supported by real user input [3].

Advantages of Heuristic Evaluation:

1) Cost-effective, requiring fewer resources compared to user testing.
2) Quick to conduct, especially with experienced evaluators.
3) Helps identify major usability issues in early design stages before software deployment.
4) Does not require participation from real users.
5) Systematic evaluation using established usability heuristics that can catch obvious errors.

Limitations of Heuristic Evaluation:

1) Results may be influenced by evaluator bias or subjectivity.
2) Lacks direct input from real users, possibly missing user-specific issues.
3) May highlight issues that are not critical or relevant to actual users.
4) Does not consider contextual or task-specific user behaviors.
5) Requires multiple experienced evaluators (typically 3–5) for more reliable results.

3. System Usability Scale (SUS)

System Usability Scale (SUS) is a popular usability evaluation method that quantifies the perceived usability of a system by the user. It was developed by John Brooke in 1986 and provides efficient and effective means of testing usability in the form of a standardized questionnaire [4]. SUS contains 10 items with each item rated on a five-point Likert scale from Strongly Disagree to Strongly Agree. Following the completion of a set of specified tasks on a website or system, participants give their ratings on statements like "I thought the system was easy to use". Their answers are subsequently translated into one score between 0 and 100, with 68 being commonly used as the average benchmark [5]. In spite of its age and simplicity, SUS is still a valid and reliable usability testing tool that is extensively used in UX evaluations and comparison of various interface designs and testing usability improvement over time.

The 10-item questionnaire, as originally developed by Brooke [4], are as follows:

#	Statement	1	2	3	4	5
1	I think that I would like to use this system frequently.	☐	☐	☐	☐	☐
2	I found the system unnecessarily complex.	☐	☐	☐	☐	☐
3	I thought the system was easy to use.	☐	☐	☐	☐	☐
4	I think that I would need the support of a technical person to use it.	☐	☐	☐	☐	☐
5	I found the various functions in this system were well integrated.	☐	☐	☐	☐	☐
6	I thought there was too much inconsistency in this system.	☐	☐	☐	☐	☐
7	I would imagine that most people would learn to use this system quickly.	☐	☐	☐	☐	☐
8	I found the system very cumbersome to use.	☐	☐	☐	☐	☐
9	I felt very confident using the system.	☐	☐	☐	☐	☐
10	I needed to learn a lot before I could get going with this system.	☐	☐	☐	☐	☐

Scale:

- 1 = Strongly Disagree
- 2 = Disagree
- 3 = Neutral
- 4 = Agree
- 5 = Strongly Agree

How to calculate the SUS score:

- Each statement with odd rating (1,3,5): subtract 1 from the user response.
- Each statement with even rating (2,4): subtract their value from 5.
- Add up the new results and multiply the total by 2.5 to get the final score (out of 100) [4].

The statements of the SUS and their order must stay the same to get reliable results and compare them with the SUS results of other websites and maintain consistency and standardized scoring. Any alteration would no longer qualify as a valid SUS test.

A recent study conducted by Ilyas et al. applied the System Usability Scale (SUS) to evaluate the usability of the Punjab Public Service Commission (PPSC) e-government website. The study collected data from 19 users categorized as novice, intermediate, and experienced, based on their familiarity with the website. The study revealed notable differences in usability perceptions, with experienced users rating the website higher (SUS score: 74), while novice users rated it significantly lower at just 19.5. Intermediate users scored the site at 55. The overall average SUS score was 62.03, placing the website in the marginal range for usability. These results highlighted how user experience level influences usability perception and underscored the value of SUS in identifying design shortcomings that affect different user groups. The study concluded that while the site was usable for returning users, improvements in interface design and clarity were necessary to support first-time visitors. This study demonstrates the effectiveness of SUS as a stand-alone method to assess website usability in real-world settings [6].

Advantages of System Usability Scale:

1) Gives you a general idea of overall usability quality.
2) Simple, fast, and cost-effective to implement.
3) Widely used and acceptable across various industries.
4) Delivers reliable, quantitative scores for easy comparison.
5) Proven to be reliable over time.
6) Platform independent, it can be used on websites, software and hardware.

Limitations of System Usability Scale:

1) Identifies the usability level but not specific issue and how to fix them.
2) Can be subjective and questionnaires may confuse some users.
3) Requires users with prior website experience.
4) Not suitable for evaluating early-stage or conceptual designs.
5) Cultural/language interpretation, it may vary across user demographics or translations.
6) It focuses on usability, not how appealing or accessible a website is.

4. Eye Tracking

Eye tracking is a usability testing method that measures and analyzes users' eye movements to determine their visual attention and interaction patterns with digital interfaces. Through the capture of data on fixations (where the gaze rests), saccades (sudden eye movements between fixations), and scan paths (the order of fixations), researchers can gain information about user behavior and cognitive processes while interacting with a system. Modern eye tracking research applies both software and hardware solutions, hardware involves wearable devices such as smart glasses and remote eye trackers, which sense movements and eye positions [7], [8].

Data collection and analysis are carried out using software platforms such as Tobii Pro Lab and OGAMA [9]. These facilities allow for the generation of visualizations such as heatmaps and gaze plots, showing regions of interest and the progression of visual attention on the interface.

Eye-tracking data are expressed via a set of visualization metrics that support the interpretation of users' visual behavior and interaction with an interface. Some of the most widely applied metrics are [7], [8]:

1) **Heatmaps**: They display visually where on a page or screen the focus of user attention lies. High visual focus points are warmer and therefore shown with warmer colors, i.e. red. Less visually prominent points look cooler and so, use cooler colors, i.e. blue or green. It gives designers the cue that it's there, when identifying areas in particular on-screen locations attracting highest visual interest.
2) **Area of Interest (AOIs)**: AOIs are previously specified interface areas that researchers chose to be scrutinized. AOIs allow collection metrics, for instance, number of fixations and fixation time, which support intense analysis of a particular portion of the interface like navigation panes, controls, and banner advertisements.
3) **Fixations and Gaze Points**: Fixations are the intervals of time when a user's eyes stay relatively stable on a given location, pointing to cognitive processing of the object. Gaze points, however, are the unprocessed eye locations at particular time intervals. Both of them assist in following the user's visual path and inferring attention and focus levels.
4) **Revisits**: This measure expresses how many times a user goes back to gaze at a location previously viewed. An abundance of revisits might indicate that the region is holding complex or attention-grabbing material, or the user couldn't process information during the first pass.
5) **Time to First Fixation (TTFF)**: The time taken by a user to first glance at a particular region of interest. This measure is used to gauge the salience or discoverability of critical UI elements.
6) **Fixation Duration**: The mean time spent on every fixation. Longer times could reflect deeper cognitive processing or struggle in comprehending a certain element.
7) **Saccade Length**: The length of fixations, and hence potentially the ability to estimate search efficiency or scanning visual behavior.

A study by Mateja applied eye tracking to assess the usability of product specification formats in online furniture store interfaces. The research involved 27 participants who viewed three different webpage designs, each displaying the same product (a corner sofa) but varying in how the specifications were presented: in plain text, tabular, and key-value formats. Using a Tobii Pro X3-120 eye tracker and iMotions software, researchers collected data on gaze patterns, dwell time, fixation counts, and user feedback. The study revealed that the key-value format, which displayed information clearly and hierarchically on the right side of the screen, was the most effective. It resulted in the shortest time to first fixation and the highest user satisfaction. Conversely, the plain text format led to the longest dwell time and lowest usability ratings, suggesting difficulty in locating information. The results demonstrated how eye tracking can

objectively uncover usability issues by highlighting where users look, how long they spend searching for content, and which formats best support quick and intuitive information retrieval. This study illustrates the practical application of eye tracking as a usability method to optimize content layout in e-commerce settings [10].

Advantages of Eye Tracking:

1) Determines where on a website the most attention is drawn.
2) Evaluates the effectiveness of visual hierarchy and placement of content.
3) Finds design errors in websites and checks if users detect important features such as buttons or messages.
4) Facilitates optimization of navigation design and minimizes cognitive load.
5) Supports A/B testing by revealing visual performance differences.
6) Beneficial for e-commerce, education, and accessibility studies.
7) Based on user's subconscious behavior instead of self-reporting prone to bias.

Limitations of Eye Tracking:

1) Expensive hardware and software requirements.
2) Requires trained personnel for setup and analysis.
3) Less effective for early-stage wireframes or low-fidelity prototypes.
4) Data can be influenced by external factors (e.g. lighting, distractions).
5) May be intrusive or uncomfortable for some participants.
6) Gives insight into attention, but not necessarily into intent or user's motivation which still need to be verbalized.
7) Users may behave unnaturally knowing that they are been monitored.

5. A/B Testing

A/B testing or split testing is a quantitative usability method for comparing two versions of an interface to ascertain which works better on predetermined behavior measures like conversion or click-through rates. The participants are allocated to two variations at random - normally a control (A) and a treatment (B) - and outcomes are statistically compared to establish if there is a difference in performance. This technique is based on the methodologies of randomized controlled experimentation and is used extensively in live settings to analyze the causal impacts of design modifications [11], [12].

Its power is in isolating individual interface components and quantifying their direct effects on user action, thus making it a credible tool for iterative interface design optimization. As Kohavi et al. highlight, the success of A/B testing relies on strict experimental design, that is, appropriate randomization, choice of metrics, and interpretation of data to provide valid results [12].

In usability applications, A/B testing facilitates evidence-based improvement of aspects like button location or content structure. It is particularly useful in busy digital sites, where small enhancements can result in significant usability and business benefits. Software like Optimizely, VWO, Adobe Target, Convert, and AB Tasty facilitates the conduct of these experiments with built-in analysis and delivery features [13].

Common metrics evaluated in A/B testing - many of which are detailed in Kohavi, Tang, and Xu (2020) [12] - include the following:

1) **Conversion Rate**: The percentage of people who carry out a particular desired action on a site such as buying or signing up.
2) **Click-Through Rate (CTR)**: Quantifies the percentage of people who click on a particular link or add following exposure (impressions). It can be used to test the efficiency of calls to action.
3) **Task Completion Rate**: Reports the ratio of users who complete a particular task to all users who were planning to.
4) **Time on Task**: Compares how long users take to complete a task. Shorter time could indicate improved usability.
5) **Bounce Rate**: Reports the ratio of users who came to a site and then exited without doing anything. Lower bounce rates are good for usability.
6) **Error Rate**: Tracks the number of errors people make when executing a given task, which can indicate problems with usability.
7) **Exit Rate**: Indicates where users are exiting a site or process, and is commonly used to measure problem areas within multi-step processes such as checkouts.
8) **Retention Rate / Return Visits**: Tracks how frequently users return, and is particularly useful in determining long-term usability and user satisfaction.

A study by Miller et al. applied A/B testing to evaluate the usability and engagement effectiveness of a clinical trial recruitment website aimed at older adults. The research was conducted as part of the STURDY clinical trial and involved 2,605 users in the first experiment and 374 users in the second, who were randomly exposed to different versions of the website landing page. Using Optimizely and Google Analytics, the researchers tested how infographic (Experiment 1) and video (Experiment 2) elements influenced user behavior. Participants were monitored for their likelihood to complete an interest form, attend a screening, or enroll in the trial. In Experiment 1, infographic Version A significantly reduced the likelihood of completing the interest form compared to the original landing page, while no significant differences were found for infographic Version B. Experiment 2 showed that one video version (Video C) significantly decreased goal conversions compared to the original, while the original version consistently produced higher engagement metrics. The findings demonstrate how A/B testing can be used to refine digital recruitment materials by empirically measuring user interaction and response, thereby optimizing usability and conversion rates in real-world web contexts [14].

Advantages of A/B Testing:

1) Provides data-driven validation of design changes by measuring real user behavior.
2) Identifies the most effective interface elements, such as layouts, buttons, and content.
3) Enhances user experience by supporting iterative design and continuous optimization.
4) Reduces risk by testing changes incrementally before full deployment.
5) Offers measurable ROI by quantifying the impact of specific design decisions.
6) Supports decision-making by combining statistical evidence with qualitative insights.
7) Delivers rapid, scalable feedback across platforms without relying on user self-reports.

Limitations of A/B Testing:

1) Requires high traffic and large samples to produce statistically reliable results.
2) Lacks diagnostic insight, revealing what works but not why it works.
3) Time-consuming to implement, with setup, monitoring, and analysis phases.
4) Less effective for early-stage designs or low-fidelity prototypes.
5) Susceptible to external variables and misinterpretation from poorly designed tests.
6) Often prioritizes short-term gains, risking neglect of long-term usability goals.
7) May deliver inconsistent experiences, as some users receive suboptimal variants.

6. Usability Testing

Usability testing is an empirical technique to assess the behavior of actual users with an interface with the objective of discovering usability issues, gathering qualitative and quantitative performance data, and determining general user satisfaction. In contrast to expert reviews, usability testing focuses on direct observation of users performing representative tasks and thus reveals friction points, misconceptions, and inefficiencies in design. In a typical usability test, test participants are given prescribed tasks to do while researchers monitor their actions, record mistakes, and occasionally pose follow-up questions to gain insight into the user's thought process. Testing may be performed in lab environments with strict control, remotely using online conferencing tools, or asynchronously using unmoderated remote techniques. Some of the most important data gathered in testing would include task completion rates, rates of error, time spent on the task, and subjective satisfaction ratings. The main benefit of usability testing is its potential to show not just whether or not users succeed or fail, but also how they work at tasks, where they get stuck, and why some design aspects can hinder success. Such findings inform iterative design refinements based on empirical evidence, not speculation. However, usability testing can be costly, involving careful planning, recruiting representative users, and systematic analysis of findings [15].

A variety of specialized tools support the usability testing process by making participant recruitment, task allocation, session recording, and data analysis easier. Some of the popular platforms are UserTesting, Lookback, Maze, PlaybookUX, and Optimal Workshop, which provide features like support for moderated and unmoderated sessions, video replays, clickstream analysis, and built-in post-test surveys [16].

Common metrics often collected during usability testing [17], include:

1) **Task Success Rate**: The percentage of participants who complete a task successfully.
2) **Time on Task**: Time participants spend on a task.
3) **Error Rate**: Number of errors occurring during task completion.
4) **Critical Incidents**: Notes of major usability failures or user annoyances.
5) **User Satisfaction Ratings**: Self-report ratings obtained via post-test survey or interview.
6) **Navigation Path Analysis**: Tracking the steps users take to complete tasks to identify inefficient or confusing routes.

A study by Subiyakto et al. applied usability testing to evaluate the design quality and user experience of an institutional repository (IR) website at a public university in Indonesia. Twelve participants were recruited based on their familiarity with the IR platform and were asked to complete five representative tasks while verbalizing their thoughts using the think-aloud protocol. The study employed screen recordings, interviews, and post-test System Usability Scale (SUS) surveys to measure usability across key metrics: task success rate (76.66%), time on task (efficiency score of 66%), and user satisfaction (SUS score of 62.3). Results revealed that users faced difficulties with search filters, unclear navigation, and misleading download links, particularly on the results and research pages. Based on these findings, the researchers proposed specific interface design improvements using Axure RP 8, including enhanced filter placement, simplified navigation menus, and clearer labeling. This study demonstrates how usability testing can effectively uncover user interaction issues and guide evidence-based website redesign in educational contexts [18].

Advantages of Usability Testing:

1) Identifies real-world problems based on actual user behavior rather than assumptions.
2) Supports iterative design by highlighting specific pain points for targeted redesign.
3) Provides both qualitative and quantitative data to inform improvements.
4) Enables discovery of unexpected user behaviors and workarounds.
5) Applicable across various product stages, from low-fidelity prototypes to live websites.
6) Facilitates understanding of user mental models and task flows.

Limitations of Usability Testing:

1) Requires substantial time and resources for planning, facilitation, and analysis.
2) Small sample sizes limit statistical generalizability of findings.

3) Participants may exhibit unnatural behavior due to observer effects (Hawthorne Effect).
4) Focuses more on detecting problems than on diagnosing underlying causes without supplementary methods.
5) Remote testing can suffer from technical issues or reduced observational richness.
6) Difficult to simulate real-world conditions perfectly in lab settings.

7. Conclusion

This paper highlights the multidisciplinary aspects of usability evaluation in website optimization. All the methods highlighted - Heuristic Evaluation, SUS, Eye Tracking, A/B Testing, and Usability Testing - offer unique insights into user behavior, system performance, and the effectiveness of the design. Though heuristic evaluation and SUS provide quick, standardized measures, eye tracking and usability testing produce deeper behavioral knowledge, and A/B testing offers empirical proof through real-time experiments. By identifying the strengths and weaknesses of each method, practitioners can more effectively match their usability plans to unique project requirements and limitations. In the end, successful website optimization relies upon the thoughtful combination and application of usability techniques to build accessible, effective, and enjoyable digital products.

References

[1] J. Nielsen, "The Theory Behind Heuristic Evaluations," Nielsen Norman Group, November 1, 1994. [Online]. Available: https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/theory-heuristic-evaluations/. [Accessed: April 15, 2025].
[2] J. Nielsen, "10 Usability Heuristics for User Interface Design," Nielsen Norman Group, Jan. 30, 2024. [Online]. Available: https://www.nngroup.com/articles/ten-usability-heuristics/. [Accessed: April 15, 2025].
[3] S. Habib, T. K. Ahsyar, M. Afdal, F. N. Salisah, and S. Syaifullah, "Enhancing Website Usability by Utilizing Heuristic Evaluation and User Feedback for Better User Experience," Journal of Information System Research (JOSH), IV (4), pp. 1093–1101, 2023.
[4] J. Brooke, "SUS: A 'quick and dirty' usability scale," in Usability Evaluation in Industry, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and A. L. McClelland, Eds. London: Taylor & Francis, 1996, pp. 189–194.
[5] [5] A. Bangor, P. Kortum, and J. Miller, "An empirical evaluation of the System Usability Scale," International Journal of Human–Computer Interaction, XXIV (6), pp. 574–594, 2008.
[6] A. Ilyas, S. H. Wajid, and A. Muhammad, "Usability evaluation of E-Government website: A use of System Usability Scale," Pakistan Journal of Engineering and Technology (PakJET), V (1), pp. 11–15, 2022.
[7] A. T. Duchowski, Eye Tracking Methodology: Theory and Practice, 3rd ed., Springer, Cham, 2017.
[8] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. van de Weijer, Eye Tracking: A Comprehensive Guide to Methods and Measures, Oxford University Press, Oxford, 2011.
[9] A. Voßkühler, V. Nordmeier, L. Kuchinke, and A. M. Jacobs, "OGAMA (Open Gaze and Mouse Analyzer): Open-source software designed to analyze eye and mouse movements in slideshow study designs," Behavior Research Methods, vol. 40, no. 4, pp. 1150–1162, Nov. 2008.
[10] A. Mateja, "Usability research of an online store using eye tracking: a comparison of product specification formats," Procedia Computer Science, CCXXV (–), pp. 3233–3242, 2023.
[11] R. Kohavi, R. Longbotham, D. Sommerfield, and R. M. Henne, "Controlled experiments on the web: survey and practical guide," Data Mining and Knowledge Discovery, XVIII (1), pp. 140–181, 2009.
[12] R. Kohavi, D. Tang, and Y. Xu, Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing, Cambridge University Press, Cambridge, 2020.
[13] ConversionXL, "25 of the Best A/B Testing Tools for 2025," conversionxl.com, Nov. 21, 2024. [Online]. Available: https://conversionxl.com/blog/ab-testing-tools/. [Accessed: May. 5, 2025].
[14] H. N. Miller, T. B. Plante, K. T. Gleason, J. Charleston, C. M. Mitchell, E. R. Miller III, L. J. Appel, and S. P. Juraschek, "A/B design testing of a clinical trial recruitment website: A pilot study to enhance the enrollment of older adults," Contemporary Clinical Trials, vol. 111, p. 106598, 2021.
[15] J. Rubin and D. Chisnell, Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests, 2nd ed., I Wiley, Indianapolis, 2008.
[16] M. Krug, "Top Usability Testing Tools to Improve Your UX Research," uxplanet.org. [Online]. Available: https://uxplanet.org/. [Accessed: May 10, 2025].
[17] T. Tullis and W. Albert, Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics, 2nd ed., Morgan Kaufmann, 2013.
[18] A. Subiyakto, Y. Rahmi, N. Kumaladewi, M. Q. Huda, N. Hasanati, and T. Haryanto, "Investigating quality of institutional repository website design using usability testing framework," International Journal of Human–Computer Interaction, XXXIX (4), pp. 299–315, 2023.