Visual Accessibility and Readability in User Interfaces: An Eye-Tracking Study

Karol Łazaruk
Department of Computer Science
Lublin University of Technology
Lublin, Poland
email: k.lazaruk@pollub.pl

Piotr Izydor Tokarski
Department of Computer Science
Lublin University of Technology
Lublin, Poland
email: p.tokarski@pollub.pl

Małgorzata Plechawska-Wójcik
Department of Computer Science
Lublin University of Technology
Lublin, Poland
email: m.plechawska@pollub.pl

***Abstract*—The aim of this work is to investigate how specific visual features of graphical interfaces influence the efficiency of information localization by users. This study focuses on examining the role of readability in visual characteristics, such as contrast, text formatting, and element highlighting, in shaping the speed and accuracy with which users can identify relevant information. The experiment involved comparing two versions of a museum interface: one designed in accordance with the principles of universal design and the second one with reduced readability. Eight stimuli were used in the study, which was conducted with 15 participants using eye-tracking technology. The qualitative and quantitative analyses were performed. Among the analysed features, task completion time, number of fixations, time to first fixation, and fixation dwell time were considered. Moreover, heat maps were also included in the study. The results indicate that high contrast, larger font size, and a well-structured content layout significantly reduce the number of fixations and shorten the time needed to locate information. Additionally, larger graphic elements with intense colours attract users' attention more effectively, and their highlighting facilitates localization. The findings are expected to contribute to a deeper understanding of user interaction with visual interfaces and to provide guidelines for designing more effective and user-friendly digital environments.**

Keywords-eye tracking; readability; visual accessibility; user interfaces.

I. INTRODUCTION

The User Interface (UI) is the main point of human-computer interaction [1]. It represents the perceptual layer, which often directly influences visual attention patterns even before cognitive processing occurs. Thus, there is a particular need to investigate how digital accessibility elements, such as visual accessibility and readability, affect user behaviour. There is also a lack of analyses that take into account how users truly perceive interfaces, rather than assessing this aspect only in terms of compliance with guidelines.

In the context of digital systems, UX refers to the user experience, which includes aspects such as usability, intuitiveness, efficiency, satisfaction, and accessibility [2][3]. The UI directly impacts the quality of the UX and, more importantly, serves as its visual and interactive element. That is why it is so important for designed interfaces to be accessible, i.e., usable by the largest possible group of users regardless of their age, background, skill level or other limitations [4], and usable, i.e., the use of the interface should be effective and easy [5].

Both usability and accessibility are the subject of numerous studies that explicitly emphasize the great importance of inclusive interface design [6][7]. Research shows that failure to comply with basic principles in this area can lead to a decline in the quality of the user experience, which also translates into, among other things, a decline in satisfaction, efficiency, and engagement [8][9]. The solution is to apply universal design principles, which define a set of guidelines that should be followed during the design process, allowing for the developed solution to be used, to the greatest extent possible, by all people, regardless of their abilities or disabilities [10].

One element of accessibility is visual accessibility, which covers the aspect of visual perception. It determines the extent to which content is presented in a readable and understandable way for the widest group of users [11][12]. Aspects of visual presentation increase visual accessibility and apply to all users [13], not just those with disabilities.

An element of usability is readability, which refers to the ease of reading content [14]. Key aspects affecting readability are contrast, as well as appropriate font type and size [15][16]. Furthermore, the appropriate structure and layout of elements increase user engagement and task completion efficiency [17][18].

Ready-made solutions in the form of heuristics [19] and accessibility standards not only facilitate interface design, but can also be used to evaluate existing solutions. Furthermore, visual accessibility and readability can be tested using automated tools, such as WAVE [20]. A tool that allows for an objective assessment of both perceptual and cognitive aspects is eye tracking [21].

The aim of this paper is to examine how different aspects of visual accessibility and readability affect the effectiveness in terms of users' ability to locate information in two versions of a museum application interface. To achieve this goal, an eye-tracking study was performed.

Unlike previous studies that primarily formulate general usability and accessibility guidelines, this study provides a quantitative validation of selected visual accessibility principles within a specific application domain, namely museum interfaces. Moreover, the research contributes by linking eye-tracking metrics (such as fixation count, dwell time, and time to first fixation) with concrete interface design variables, including contrast, typography, and layout structure. This allows not only confirmation of known principles, but also their measurable impact on user attention and search efficiency in realistic interaction scenarios.

The remainder of this paper is structured as follows. In Section II, the materials and methods used in the study are described. Section III presents the results of the eye-tracking analysis. Section IV discusses the findings and their implications. Finally, Section V concludes the paper and outlines directions for future research.

II. Materials and Methods

Fifteen participants (14 males and 1 female) took part in the study, with an average age of 23.2 years (SD = 0.54). Prior to the study, informed consent was obtained from all participants (the opinion from the Scientific Research Ethics Committee of the Lublin University of Technology No. 1/2024 of 19 February 2024). All participants had experience in using websites.

The research object was a proprietary online museum website created using the Angular library. Two alternative versions of the interface were implemented: the first, designed based on universal design principles (Interface 1), and the second, in which no good practices were applied and readability was intentionally limited (Interface 2).

Eye movements were recorded using a Gazepoint GP3 HD eye tracker (Gazepoint, Vancouver, Canada; 150 Hz, accuracy 0.5-1°) on a 17.3" screen (1920 × 1080 pixels).

The experimental procedure involved a short introduction to the experiment, in which the participants were informed about the aim and procedures. The actual eye tracking study was conducted on a group of participants who were asked to complete a total of 13 tasks (commands): five tasks (1-5) in two versions of the application interface, and three additional commands unrelated to any version of the interface (6-8), which were utilised for additional verification of the readability aspect. Table 1 presents an overview of the tasks performed by the respondents, along with the visual accessibility and readability issues that were the subject of their assessment. To minimize potential learning effects associated with the repeated exposure to similar interface structures, the order of interface presentation was counterbalanced across participants. Half of the participants started with Interface 1, while the other half began with Interface 2. Additionally, the order of tasks was randomized within each interface version. This approach was applied to reduce bias related to familiarity and learning effects in the paired experimental design.

TABLE I. Research Tasks with Corresponding Issues of Visual Accessibility and Readability

Task number	Task scope	Area of accessibility/readability
1	Locating the link to the museum's social media	Visual contrast of the interface
2	Locating the question in the FAQ section	Aesthetics, formatting, and grouping of text content
3	Locating the link to the latest exhibition	Organisation and order of elements
4	Locating the button/link to the museum's rules and regulations	Distinguishability of buttons and links
5	Locating the link to the FAQ section	Visibility and location
6-8	Locating the image	Readability of typography (font type, colour, size), visibility of visual elements against the background, relationship between captions and graphic elements, and visual hierarchy of elements

Each task was associated with a particular visual stimulus in the form of a screenshot of the application and required participants to localise specific interface elements that were formulated as commands (e.g., 'Locate the museum's social media button'). The participants could move on to the next tasks by pressing the space bar. The tasks were displayed one at a time.

Heat maps were exported for the purpose of qualitative analysis. The heat maps allow for a visual representation of the aggregated level of visual attention, which is represented by colours applied to the visual stimulus. This makes it possible to locate areas that attract attention as well as those that are ignored by users [22].

An Area of Interest (AOI) was defined for the collected eye tracking data. The AOI data was used in qualitative analysis, which was conducted on the basis of proprietary scripts in the R language. The following eye tracking metrics were selected and considered in the context of the defined AOIs:

- Number of fixations - the total number of fixations in the AOI, which is negatively correlated with the performance [23].
- Time To First Fixation (TTFF) - the time elapsed from the appearance of the stimulus to the first fixation in the AOI. This measure indicates how quickly a particular element attracts the user's attention - the shorter the time, the faster the observer's attention is attracted [24].
- Fixation dwell time - the total time that the user spent in the AOI. It indicates the level of interest - the longer the time, the greater the interest [25].

An additional metric was the task completion time, which refers to the time taken to complete a particular task. It measures the performance of completed tasks - shorter times indicate that specific tasks are easier to complete and that the interface design is better [26].

Normality and homogeneity of variance tests were performed within the quantitative eye-tracking data to verify the applicability of Student's t-test. As the assumptions were not met, a non-parametric equivalent of Student's t-test, i.e., Wilcoxon signed-rank (paired) test, was performed.

III. RESULTS

Based on eye-tracking data, the heatmaps were generated. Their content made it possible to verify that, for tasks 1-5 in Interface 1, areas receiving a high level of attention covered a smaller region and were significantly fewer in number compared to the corresponding heat maps for Interface 2.

Figures 1 and 2 illustrate sample heat maps associated with task 2, which involved finding a specific question in the FAQ section. In Figure 1, it can be observed that in Interface 1, the participants' gaze was concentrated most intensely on the element representing the target solution of the task, but it is also worth noting the increased intensity of attention within another section of the FAQ, where the participants focused their gaze when reading the content of the preceding section. Several areas of low attention intensity within the navigation bar and other questions can be seen, but their intensity is significantly lower. In the case of Figure 2, it can be observed that the area of increased visual attention has a much larger surface area than in Interface 1, where attention is concentrated on only two relatively small, dense points. Furthermore, this area has a generally higher level of attention intensity, indicating that in Interface 2, the subjects had to put more effort into finding the target element.

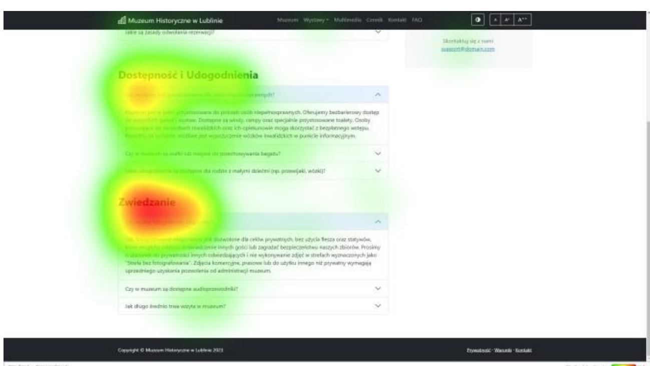

Figure 1. Example of a heatmap of task 2 in the case of Interface 1.

Figure 2. Example of a heatmap of task 2 in the case of Interface 2.

In the case of three additional commands (6-8), heat maps clearly show that the visual attention of the participants was focused primarily on the highlighted elements. In the case of graphics, these were large objects in bright colours, while the text was visually appealing due to its clear, sans-serif font in bright but well-contrasting colours. Similarly, smaller elements were more difficult to locate, resulting in a greater number of areas of moderate or increased attention intensity.

Figure 3 shows a heat map associated with task 8, which shows that the large highlighted element with a bright contrasting font attracts the most attention compared to other less prominent elements.

Figure 3. Example of a heatmap of task 8.

For the purposes of qualitative analysis, mean values, medians, and confidence intervals were calculated for the individual tasks in both interfaces: task completion time and (in the context of AOI) fixation dwell time, TTFF, and number of fixations.

Figure 4 presents a comparison of mean task completion times for individual tasks for both interfaces. As can be easily seen, the mean times for Interface 1 were shorter than those for Interface 2 in each command. The difference between the times ranged from 0.6 s to approximately 7.0 s on average.

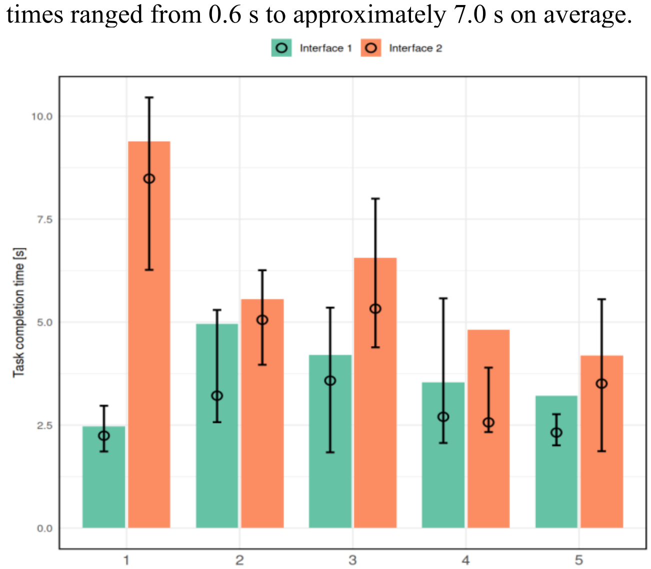

Figure 4. Mean task completion time per task according to the interface obtained for all participants.

Figure 5 presents a comparison of mean fixation dwell times (AOI) in individual tasks for both interfaces. Similar to the mean task completion time, the mean dwell times for Interface 1 were also shorter than those for Interface 2 for each command. Only for task 2, the mean results are almost identical, while in the other commands, the differences between the times ranged up to approximately 1.6 s on average.

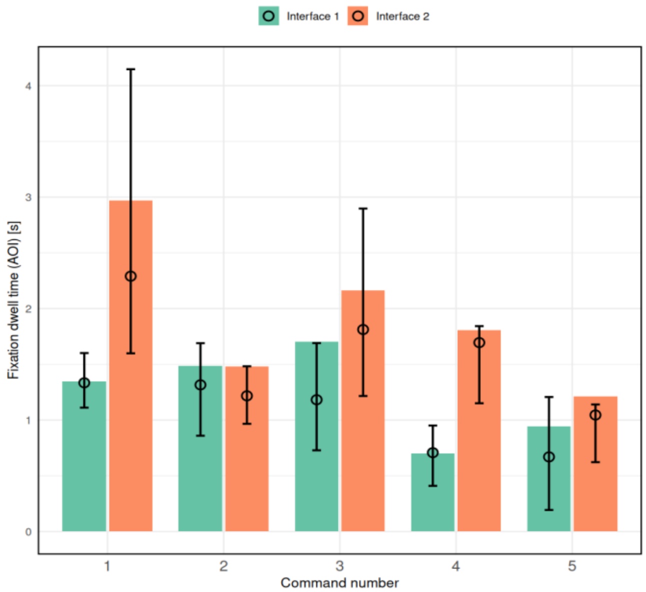

Figure 5. Mean fixation dwell time (AOI) per task according to the interface obtained for all participants.

In the case of the TTFF (AOI) measure, for which the mean values for individual commands are presented in Figure 6, it can be seen that the TTFF time for tasks 2, 3, and 5 was greater in the case of Interface 2, averaging approximately 0.5 s to 1.5 s. For task 1, the TTFF times are similar, while for task 4, the discrepancy is clearly noticeable and is approximately 0.9 s less for Interface 2.

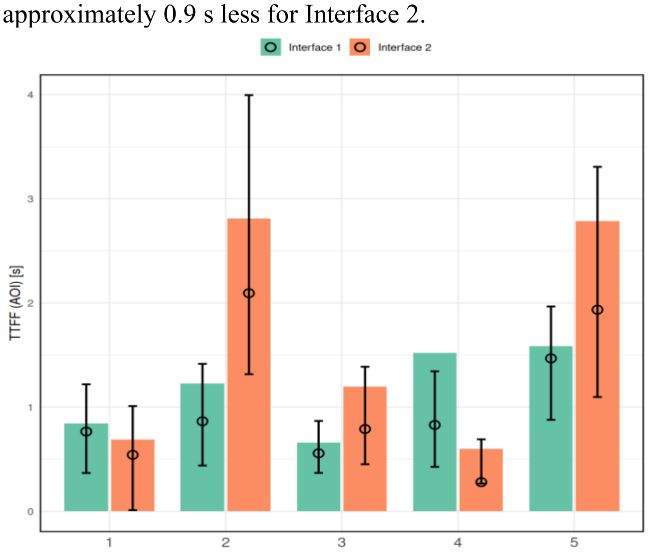

Figure 6. Mean TTFF (AOI) per task according to the interface obtained for all participants.

A comparison between the mean fixation counts for both interfaces across all commands presented in Figure 7 demonstrates that, except for task 5, the mean fixation count was higher, ranging from approximately 4.4 to 11 fixations for Interface 2. In the case of command 5, this number was very close for both interfaces.

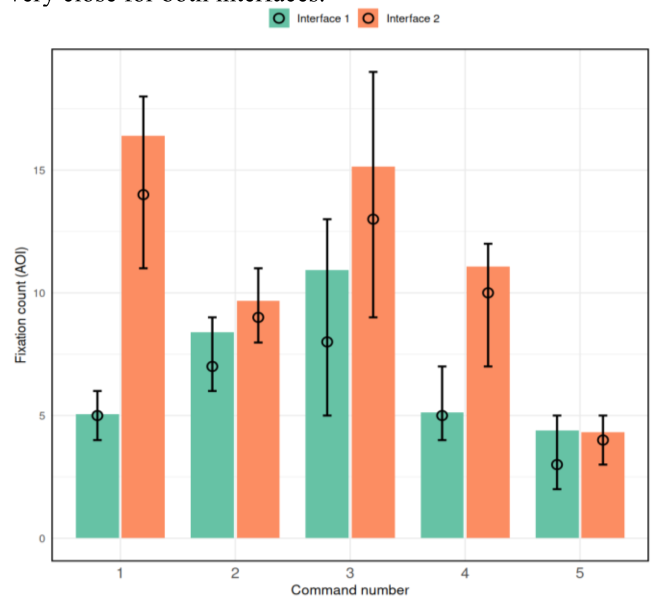

Figure 7. Mean number of fixations (AOI) per task according to the interface obtained for all participants.

Normality and variance heterogeneity tests were performed for every command for each measure for both interfaces. The test results indicated that for some data, the assumptions of the Student’s t-test were not met, so non-parametric Wilcoxon signed-rank (paired) tests were performed. The tests showed that there are statistically significant differences between interfaces that are not applicable to some commands. In the case of the task completion time, the statistically significant differences were observed for command 1, and similarly for dwell time for commands 1 and 4. Regarding TTFF, a significant difference between interfaces was observed for command 2. In the case of fixation count, the differences were observed for commands 1, 3 and 4. All values met the criterion of $p < 0.05$.

IV. DISCUSSION

As can be observed from the heat maps, participants’ attention is more clearly focused on key areas in Interface 1 than in Interface 2. These areas are smaller, yet they attract more visual attention. In Interface 2, however, participants’ attention is less concentrated, which is reflected by a more even distribution of attention and the presence of areas with lower fixation intensity. This suggests that, even without prior experience with the application, users can intuitively locate the most important functions in Interface 1. The results also confirm the conclusions presented in [17], which state that elements located in the top, middle, and right sections of the interface attract the most attention.

Similar observations can be made regarding heat maps relating to contrast, font size, the order of elements, and the readability of navigational elements. Participants were able to use the website more easily in high contrast mode, and elements with larger font sizes were more visible. Furthermore, Interface 1, with clearly organised and logically ordered elements, was associated with higher levels of visual attention. Key navigational components, such as clearly visible FAQ buttons in the navigation bar, attracted more attention than less prominent links in the footer. These observations are consistent with the results presented in article [16], which found that participants preferred text written in larger font sizes, as well as high contrast between the font and background colours.

In the case of heat maps related to visual aspects and interface readability, such as layout, element highlighting, and typography, it has been proven that larger, emphasized or intensely coloured images attract the user's attention much more than smaller, less prominent, or less noticeable elements. Heat maps suggest that the participants looked at smaller areas, indicating greater intuitiveness and simplicity of Interface 1. These results are consistent with the conclusions of the study [15], in which the authors demonstrated that bright, saturated colours effectively attract users' attention. Furthermore, a study conducted in [18] shows that large images, photos of celebrities, minimal text, and search functions attract attention, and this study proves the same with regard to large images and minimal text.

In the context of quantitative analysis, for both task completion time and fixation dwell time, it can be observed that Interface 1 was characterised by shorter values compared to Interface 2. This leads to the conclusion that increasing visual accessibility and readability significantly improves the efficiency with which the participants are able to find specific elements within the interface. However, it is worth noting that the differences in fixation dwell time between the interfaces were smaller than the differences in task completion time. This is because fixation dwell time is only part of the total task completion time within the AOI and does not include, for example, the time needed for navigation, information processing, and decision-making by the participants.

It should also be noted that in task 2, the fixation dwell times were very similar, which may be due to the fact that in both interfaces, the subjects had to analyse the same text fragments, and the only difference was in the aesthetics, formatting, and grouping of the text content. Therefore, these visual changes did not significantly affect fixation time, although they still facilitated orientation in the content and influenced the overall efficiency of task performance, which is clearly noticeable in the other measures (task completion time, TTFF, and number of fixations).

The TTFF measure revealed that the times associated with tasks 2, 3, and 5 were notably faster for Interface 1 compared to Interface 2. This suggests that Interface 1 has superior visibility, aesthetics, and content organisation, enhancing the perceptibility of its key areas. Furthermore, the shorter TTFF time suggests that Interface 1 is more effective at managing user attention, is more readable, and is visually clear.

By contrast, it is worth noting the opposite trend for tasks 1 and 4, for which TTFF times were shorter for Interface 2. This is because, in both cases, the page content included a map that took up most of the screen and attracted participants' attention. In case of Interface 1, the map was placed above the target text in task 1, making it slower to locate the necessary information than in Interface 2, where the map was located below the target text. In case of Interface 1 in task 4, only the text content was displayed for the participants to analyse. These results suggest that compliance with universal design principles does not always guarantee shorter TTFF, and that the layout and nature of page content also significantly impact the effectiveness of information retrieval.

The mean number of fixations confirmed that in almost all cases, the number of fixations for Interface 1 was lower than for Interface 2. This suggests that better organisation and readability of interface elements, as well as contrast, content aesthetics, and the use of buttons instead of links, contribute to more effective content searching.

However, it is worth noting that, in task 5, locating the FAQ link in Interface 1 took less time than in Interface 2 but required almost the same number of fixations. This suggests that the greater readability and better placement of the element (the link in the navbar with the larger font) enabled it to be identified more quickly, despite the similar number of fixations. In Interface 2, however, the link in the footer with the smaller font required a similar number of fixations but took more time to be located.

V. Conclusion and Future Works

The aim of the study was to assess the impact of visual accessibility and interface readability on the effectiveness of information location by users, based on a comparison of two versions of a museum web application. Eye-tracking and task-related metrics were utilised to perform the analysis. The study contributes to the field by providing quantitative evidence of how specific visual design features influence user attention and performance within a domain-specific interface.

The study indicates that users are more likely to notice graphic elements that are larger, have intense colours, or are highlighted or emphasised. Additionally, visual interface features, such as high contrast, larger font size, aesthetic formatting, and logical content organisation significantly improve visual accessibility and reduce the time needed to find information. These results demonstrate that an effective layout and visual distinction of elements can enhance the efficiency of navigating the interface and provide valuable insights for creating more user-friendly digital environments.

While the study provides relevant findings and clear conclusions, it should be noted that it has certain limitations. First, the relatively small sample size may limit the generalizability of the results. Non-parametric Wilcoxon signed-rank (paired) tests revealed no statistically significant differences between the two interfaces in relation to certain assigned tasks and measures. Future research plans to examine a significantly larger group of respondents and expand the number of visual stimuli to investigate more aspects that could affect visual accessibility and readability. Second, the within-subject (paired) experimental design introduces a potential risk of learning effects, although counterbalancing and task randomization were applied to mitigate this issue. Third, the study relies solely on eye-tracking data, without incorporating predictive models, such as visual saliency analysis. An important extension of the presented study could involve the use of computational visual saliency models. Such models allow prediction of areas that are likely to attract user attention prior to actual interaction. Comparing predicted saliency maps with empirical eye-tracking data could provide deeper insight into discrepancies between expected and observed gaze behavior. This would further strengthen the evaluation of interface effectiveness and support more predictive design approaches. Future work should address these limitations by increasing the sample size, integrating computational attention models, and exploring additional interface types and domains.

REFERENCES

[1] A. Pitale and A. Bhumgara, “Human computer interaction strategies—designing the user interface,” Proc. 2019 Int. Conf. Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2019, pp. 752–758, doi:10.1109/ICSSIT46314.2019.8987819.

[2] O. Badran and S. Al-Haddad, “The impact of software user experience on customer satisfaction,” J. Management Information and Decision Science, vol. 21, pp. 1–20, 2018.

[3] M. Hassenzahl and N. Tractinsky, “User experience—A research agenda,” Behaviour & Information Technology, vol. 25, pp. 91–97, 2006, doi:10.1080/01449290500330331.

[4] O. Sohaib, W. Hussain, and M. Khalid, “User experience (UX) and the web accessibility standards,” Int. J. Computer Science Issues, vol. 8, pp. 584–609, 2011.

[5] M. Matera, F. Rizzo, and G. Carughi, “Web usability: principles and evaluation methods,” in Web Engineering, 2006, pp. 143–180, doi:10.1007/3-540-28218-1_5.

[6] J. Gómez-Delgado, C. Marín-Palacios, and L. Moreno, “Integrating accessibility, usability and UX in web project management: a systematic review,” Univ. Access Inf. Soc., vol. 25, p. 28, in press, doi:10.1007/s10209-025-01298-0.

[7] V. Nasr, M. Benden, and M. Zahabi, “An approach for assessing usability and accessibility of assistive technology for persons with disabilities,” Applied Ergonomics, vol. 129, art. no. 104581, 2025, doi:10.1016/j.apergo.2025.104581.

[8] T. Tullis and W. Albert, Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics, 2nd ed. Burlington, MA: Morgan Kaufmann, 2008.

[9] M. Ekin et al., “Impact of web accessibility on cognitive engagement in individuals without disabilities: evidence from a psychophysiological study,” PLoS One, vol. 20, no. 7, art. e0328552, July 2025, doi:10.1371/journal.pone.0328552.

[10] J. Lazar, D. Goldstein, and A. Taylor, “Evaluation methods and measurement,” in Ensuring Digital Accessibility Through Process and Policy, San Francisco, CA: Morgan Kaufmann, 2015, pp. 139–159.

[11] G. E. Legge, “Visual accessibility: a challenge for low-vision research,” Optometry and Vision Science, vol. 91, no. 7, pp. 696–706, July 2014, doi:10.1097/OPX.0000000000000310.

[12] K. Cornish, J. Goodman-Deane, K. Ruggeri, and P. J. Clarkson, “Visual accessibility in graphic design: a client–designer communication failure,” Design Studies, vol. 40, pp. 176–195, 2015, doi:10.1016/j.destud.2015.07.003.

[13] H. Petrie, F. Hamilton, and N. King, “Tension, what tension? website accessibility and visual design,” Proc. 2004 Int. Cross-Disciplinary Workshop on Web Accessibility (W4A ’04), New York, NY, USA: ACM, 2004, pp. 13–18, doi:10.1145/990657.990660.

[14] A. Miniukovich, A. De Angeli, S. Sulpizio, and P. Venuti, “Design guidelines for web readability,” Proc. 2017 Conf. Human Factors in Computing Systems, 2017, pp. 285–296, doi:10.1145/3064663.3064711.

[15] A. Lewandowska, A. Olejnik-Krugly, J. Jankowski, and M. Dziśko, “Subjective and objective user behavior disparity: towards balanced visual design and color adjustment,” Sensors, vol. 21, pp. 1–18, 2021, doi:10.3390/s21248502.

[16] L. Rello and M.-C. Marcos, “An eye tracking study on text customization for user performance and preference,” Proc. 8th Latin American Web Congress, 2012, pp. 64–70, doi:10.1109/LA-WEB.2012.13.

[17] M. Țichindelean, M. T. Țichindelean, I. Cetină, and G. Orzan, “A comparative eye tracking study of usability—towards sustainable web design,” Sustainability, vol. 13, pp. 1–31, 2021, doi:10.3390/su131810415.

[18] S. Djamasbi, M. Siegel, and T. Tullis, “Generation Y, web design, and eye tracking,” Int. J. Human-Computer Studies, vol. 68, pp. 307–323, 2010, doi:10.1016/j.ijhcs.2009.12.006.

[19] R. da Silva de Queiroz Pierre, “Heuristics in Design: A Literature Review,” Procedia Manufacturing, vol. 3, pp. 6571–6578, 2015, doi: 10.1016/j.promfg.2015.07.961.

[20] D. Mothy, A. P. Reddy, C. W. Cai, H. S. Choudhry, and M. H. Dastjerdi, “Assessing the readability, quality, and visual accessibility of patient education websites for laser refractive surgery,” Ophthalmic Epidemiology, vol. 32, no. 6, pp. 704–710, 2025, doi:10.1080/09286586.2025.2500014.

[21] S. Eraslan, Y. Yesilada, and S. Harper, “Eye tracking scanpath analysis techniques on web pages: a survey, evaluation and comparison,” J. Eye Movement Research, vol. 9, pp. 1–19, 2016, doi:10.16910/jemr.9.1.2.

[22] O. Špakov and D. Miniotas, “Visualization of eye gaze data using heat maps,” Elektronika ir Elektrotechnika, vol. 115, pp. 55–58, 2007.

[23] M. M. Porras, C. A. N. K. Campen, J. J. González-Rosa, F. L. Sánchez-Fernández, and J. I. N. Guzmán, “Eye tracking study in children to assess mental calculation and eye movements,” Scientific Reports, vol. 14, art. no. 18901, Aug. 2024, doi:10.1038/s41598-024-69800-x.

[24] P. Wlekły, M. Nermend, M. Gryczka, M. Borawski, and H. Shabani, “Use of eye-tracking in the design and optimisation of training materials for future drivers,” in Proc. 28th European Conference on Artificial Intelligence (ECAI 2025), 2025, pp. 83–96, doi:10.18276/978-83-8419-053-1-6.

[25] A. Andrychowicz-Trojanowska, “Basic terminology of eye-tracking research,” Applied Linguistics Papers, vol. 2, no. 2, pp. 123–132, 2018, doi: 10.32612/uw.25449354.2018.2.pp.123-132.

[26] Q. Wang, S. Yang, M. Liu, Z. Cao, and Q. Ma, “An eye-tracking study of website complexity from cognitive load perspective,” Decision Support Systems, vol. 62, pp. 1–10, 2014, doi:10.1016/j.dss.2014.02.007.