The role of mental workload in determining the relation between website complexity and usability: an eye-tracking study

Giovanni Serra[1], Federica De Falco[1], Piero Maggi[1], Rita Forsi[2], Antonello Cocco[2], Giancarlo Gaudino[2], Massimo Amendola[2], & Francesco Di Nocera[1]
[1]Department of Psychology, Sapienza University of Rome
[2]Higher Institute of Communications and Information Technologies, Ministry of the Economic Development
Italy

Abstract

Digital technology is now crucial for carrying out many activities and the drive for innovation in digitization has involved not only private organizations but also public administrations. However, despite a growing awareness of the importance of digitization of public services, usability issues have been addressed only recently. Terms like "Usability", "User Experience", and "Human-Centred Design" are surely becoming part of the vocabulary, but often they are used as empty keywords by policymakers, while there is a lack of specific research in this domain. One notable example is the lack of studies on cognitive load imposed by information abundant websites and its influence on both users' performance and perceptions of usability. The main objective of the present study is to evaluate the mental workload of users navigating websites with different levels of complexity in their information architecture. Eye movements of twenty users were recorded during the execution of search tasks carried out on websites of three public agencies previously selected for their complexity level. Mental workload assessment was obtained by analysing subjects' ocular behaviour. Results showed that low complexity websites were associated with better performance, lower mental workload and higher usability rates compared to high complexity websites.

Introduction

According to the European Commission (EC) "in less than a decade, most economic activities will be managed through digital ecosystems" making the digitization of all economic sectors mandatory. The European Commission has also published an "Action Plan" to guide Governments in this transition. The document establishes some basic principles to regulate and promote the digitization of Public Administrations (PAs). The Italian Government, through an internal policy document (Directive for Public Administration and Innovation of November 26, 2009, No. 8), required PAs to improve the quality and usability of their websites by applying guidelines inspired by "Usability principles". In January 2017 the Italian Digitalization Agency (Agenzia per l'Italia Digitale: AgID) published the first

In D. de Waard, K. Brookhuis, D. Coelho, S. Fairclough, D. Manzey, A. Naumann, L. Onnasch, S. Röttger, A. Toffetti, and R. Wiczorek (Eds.) (2019). Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2018 Annual Conference. ISSN 2333-4959 (online). Available from http://hfes-europe.org

version of a web toolkit indicating style sheets, Javascript components and HTML examples to be applied for easily adhering to usability standards.

Usability has been defined by the ISO 9241-11 (updated by ISO 9241-210) as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use". Effectiveness refers to the completeness and accuracy in the achievement of the objectives by the users, efficiency refers instead to the optimization of the user's cognitive and temporal resources expenditure and satisfaction concerns the aspects of comfort and acceptability of the use of the product. Interface usability is closely related to another concept defined "Information Architecture" (IA), describing the structural planning of information contained within a website (Garrett, 2010, Rosenfeld & Morville, 2006). Several studies (Gullikson et al., 1999; McGills & Toms, 2001; Miller & Remington, 2004; Rosenfeld & Morville, 2006; Toms, 2002) found that highly complex information architecture is related to decreases in perceived usability. The ease of use of an interface directly affects the cognitive effort required for the interaction. Although the relationship between usability and cognitive effort can be easily deducted, the two concepts have been studied independently. Usability has been studied applying the User-Centred Design methodology with the objective to improve the Human-Machine Interaction. Instead, the construct of mental workload (MWL) has been investigated in the Human Factors research tradition to study human error and to increase systems' security. The methodology used to measure interfaces' usability and MWL are different but share a cognitive-behavioural approach. Eye movements have been used to study visual search strategy and have been applied both to usability studies and to measure the cognitive effort imposed by a complex task (May et al., 1990; Jacob & Karn, 2003; Pan et al., 2004; Poole & Ball, 2006; Majaranta & Bulling, 2014).

Recently, the availability of less intrusive eye-tracking systems allowed researchers to effectively use indices of ocular activity as a measure of the operator mental workload (see Van Orden et al., 2001). For example, frequency and duration of eye-blinks have been found to be inversely correlated to mental load (Brooking, Wilson, & Swain, 1996; Hankins & Wilson, 1998). Additionally, some studies (Bunecke, 1987; Ephrath et al., 1980) have shown that workload affects the duration of fixations, whereas others (Bellenkes, Wickens, & Kramer, 1997; Miller, 1973) recorded shorter and more frequent fixations in expert operators.

Studies from our research group (see for example Di Nocera, Camilli, & Terenzi, 2007 and Di Nocera, Ranvaud, & Pasquali, 2015) have introduced the use of an algorithm for the analysis of the scanpath called Nearest Neighbour Index (NNI). Using this algorithm, it is possible to compare different fixation patterns during the execution of a specific task. The result is expressed by a single value varying with fixation clustering-spreading. Also, as described by Camilli et al. (2008), the index is sensitive to the type of demand, and values increase as mental workload increases when the taskload is due to the temporal demand, while they decrease as mental workload increases when the taskload is due to the visuo-spatial demand.

Eye-tracking techniques have also been applied in usability studies (see Goldberg, 2003). Optimization of the interface features and functions has been studied both

using metrics based on scanpath elements (such as fixations and saccades) and metrics based on the analysis of the entire visual exploration path. Many authors (Fitts et al., 1950; Goldberg et al., 2002; Cowen et al., 2002; Ehmke & Wilson, 2007; Poole et al., 2005; Goldberg & Kotval, 1999; Kotval & Goldberg, 1998; Ehmke & Wilson, 2007; Habuchi et al., 2008; Wang et al., 2018) have observed that an incorrect arrangement of the interface elements, leading to poor usability, is associated with a greater number and of saccades and longer duration, as well as with visual search strategies characterised by transitions between non-contiguous areas of interest.

It is important to underline that the choice and interpretation of those metrics must not be rigid. In fact, it always depends on the researcher's objectives: if the main goal of the study is to understand the effectiveness of a banner ad in capturing customers attention, a great number of fixations on it (and long durations) will generally be considered as a positive fact; on the contrary, if the purpose of the study is to investigate how easy or difficult it is to find a specific element on a web page, a great number of fixations may be associated with poor usability, as the website is likely to be more complex and difficult to understand and browse.

The sequence of eye movements occurring while the user explores an interface may provide valuable information about her browsing experience and cognitive load. Behavioural metrics allow to objectively investigate user strategies and behaviours, avoiding potential distortions that are often related to the use of subjective metrics (such as self-report questionnaires and qualitative interviews).

With that in mind, considering the importance of usability and workload in optimizing the performance of individuals interacting with an interface, a clarification of the relationship between these constructs is strongly needed. Recent studies have made an attempt in this direction, however they showed mixed results. In a recent study Kokini and colleagues (2012) found that an increase in mental workload negatively affected perceived usability. Furthermore, Fedele and colleagues (2017) have found that positive interaction experiences were associated with lower mental workload. Differently, Longo and Dondio (2015) reported no relationship at all between the two constructs and suggested that they should therefore be considered separately. Nevertheless, resources allocation is an issue in any type of interaction, and the integration of mental workload assessment into usability evaluations could have important implications for improving interface design as well as for better defining the usability construct.

The study

The present study was carried out as part of an ongoing collaboration between the Italian ICT government agency (Istituto Superiore per le Comunicazioni e le Tecnologie dell'Informazione: ISCOM) and the Eye-Tracking Laboratory at the Department of Psychology, Sapienza University of Rome. The main objective of the project, named "WEBLOAD", was to evaluate the mental workload associated with the navigation of PAs websites with different levels of complexity and understanding its effects on perceived usability.

The main hypothesis of this study was that a greater complexity of the information architecture structure would be related to higher mental workload and poor usability evaluations. Three large PAs websites (whose identity we are not allowed to disclose) were selected after a heuristic evaluation of their IA structures aimed at exploring the number of menu levels and their related categories. All websites were similar in the design and interaction features (menu, colours, aesthetic) but different in terms of information architecture complexity. From the less complex to the more complex, the identified websites will be referred as website 1, website 2, website 3 hereinafter.

Table 1: Information Architecture of the selected websites.

Website	Information Architecture structure (Number of levels and number of categories per level)							
	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	**Total**	**Complexity**
1	5	44	134	170	-	-	353	**Low**
2	5	53	160	190	53	8	469	**Medium**
3	6	45	109	266	158	56	630	**High**

Participants

Twenty volunteers participated in the study (7 females, average age = 57; SD = 6). All of them were native Italian speakers, were naïve as to the aims, the expected outcomes, and the methodology of the experiment, and had normal or corrected-to-normal vision. All the subjects declared to use the Internet every day. This study was performed in accordance with the Helsinki Declaration of the World Medical Association.

Materials and method

The X2-30 eye-tracking system (Tobii, Sweden) was used to record eye movements during the interaction with the three websites. This is a standalone eye tracker that can be used in various setup by attaching it to monitors, laptops or for performing eye-tracking on physical objects with a sampling rate of 30 Hz.

Subjective measure of perceived usability and mental workload were collected at the end of the interaction with each website using the following scales.

- Net Promoter Score (NPS: Reichheld, 2003, 2006): a single item “How likely would you recommend this website/brand to a friend or colleague?”. Subjects can answer using an 11-point scale (0 to 10). This questionnaire is based on the fundamental concept that a user (customer) falls in one of the following three categories: Promoter (provides a score between 9 and 10), Neutral (provides a score between 7 and 8) and Detractor (provides a score between 0 and 6). The final value of the NPS is obtained by subtracting the percentage of detractors from the percentage of promoters. Sauro and Lewis (2010) found a strong correlation between perceived usability and NPS. Word of mouth is critical for

the success of PAs websites and digital services. Therefore, the NPS was included as a measure.

- *Usability Evaluation 2.0:* (Us.E. 2.0: Di Nocera, 2013) a multidimensional questionnaire for the evaluation of website usability. The questionnaire consists of 19 items subdivided into three subscales representing the framework users would adopt for evaluating the quality of their interaction with the interface: (Mental) Handling, Satisfaction and Attractiveness. Us.E. 2.0 allows fast assessment of website perceived usability identifying critical issues that could be eventually addressed by more extensive testing and re-design. Users are required to answer to all items along with a 5-point Likert scale (ranging from "strongly agree" to "strongly disagree"). The "(Mental) Handling" scale measures the interaction with the structure of the website (e.g. information architecture, layout). This is the "hard" usability aspect. Low scores in this scale would suggest the need to make changes to information architecture or page layout. The "Satisfaction" scale measures the perceived utility of the website. Low scores in this scale may indicate that the website does not meet users' needs either because the users are not those expected by who created the website or because contents/services are not those expected by the users. Finally, the "Attractiveness" scale measures the interaction with the aesthetic features of the website (i.e. apparent usability). Low scores in this scale would suggest the need for a restyling.
- NASA Task Load Index (NASA-TLX: Hart & Staveland, 1988) has been administered at the end of a task. The respondent provides an evaluation of his/her perceive workload along six scales (from 0 to 100): mental demand, physical demand, temporal demand, effort, performance, and frustration. This self-report questionnaire is one of the most used to evaluate the mental workload experienced by individuals during the execution of a task (Wickens, Hollands, Banbury & Parasuraman, 2013; Hart, 2006).

Procedure

Five equivalent research tasks have been designed for each website. Subjects had to search specific information in different areas of the websites (e.g. "Transparent Administration", "Open Data", "Downloads").

Participants performed the entire test in three separate sessions of the approximate duration of 50 minutes and at about 15 days apart from each other in order to limit effects related to fatigue and task duration. Moreover, with the aim of avoiding effects related to the order of presentation of the stimuli, the websites and the tasks were randomly assigned to the participants. Prior starting each navigation session, the experimenter asked participants if they had ever interacted with the website under testing. Participants had a free navigation session to familiarize with the websites before the calibration of the eye-tracker. Therefore, they were instructed on the activity and were asked to perform the task to their best.

At the end of each task, participants also reported their perceptions about the level of complexity of the task on a scale from 1 to 5 (1 = Not difficult at all; 5 = Extremely difficult). After completing all the five tasks participants reported their personal

information (gender, age, educational qualification, employment, the frequency of internet use) using an on-line form, and answered to the satisfaction, perceived usability and mental workload scales.

Data analysis and results

Success rate, completion time, perceived complexity, NPS, Us.E. 2.0 (Handling, Satisfaction, Attractiveness), NASA-TLX, and NNI scores were analysed in repeated-measure ANOVA designs using Complexity (website 1 vs. website 2 vs. website 3) as repeated factor.

Success rate was significantly different between websites ($F_{2,36} = 3.04$; $p < .05$). Duncan post-hoc testing showed that success rate for the high-complexity website (website 3) was significantly lower than the other two.

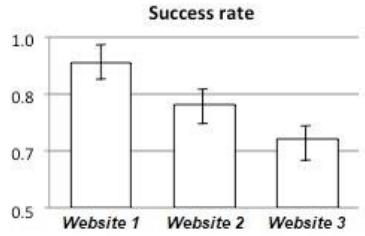

Figure 1. Success rate per website.

Completion time was significantly different between websites ($F_{2,36} = 4.02$; $p < .05$). Duncan post-hoc testing showed that completion time for the low-complexity website (website 1) was significantly faster than the other two.

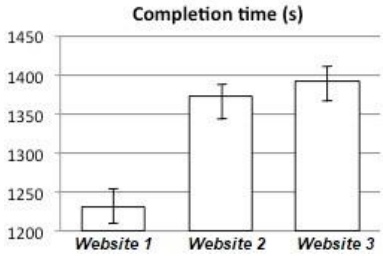

Figure 2. Completion time (average) per website.

Perceived complexity was significantly different between websites ($F_{2,36} = 3.92$; $p < .05$). Duncan post-hoc testing showed that perceived complexity of the low-complexity website (website 1) was significantly lower than the other two.

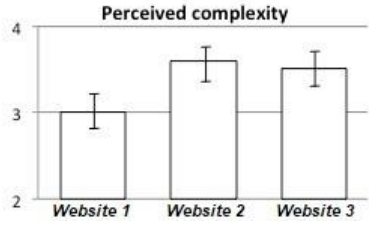

Figure 3. Perceived complexity per website.

NPS score was significantly different between websites ($F_{2,36}$ = 4.52; p < .05). Duncan post-hoc testing showed that proportion of the low-complexity website (website 1) was significantly higher than the other two.

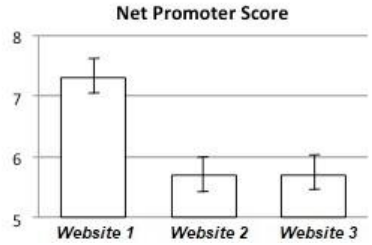

Figure 4. Net Promoter Score per website.

Us.E. 2.0 scores were significantly different between websites ($F_{2,36}$ = 5.12, p < .05) and between scales ($F_{2,36}$ = 17.77, p < .01). No significant interaction between website and scale was found. Duncan post-hoc testing showed that 1) usability for the low-complexity website (website 1) was significantly higher than the other two and 2) handling was significantly lower than satisfaction and both were significantly lower than attractiveness thus indicating (Mental) Handling as the most awkward dimension.

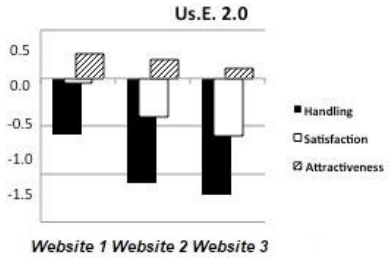

Figure 5. Us.E. 2.0 per website.

NASA-TLX score was significantly different between websites ($F_{2,36}$ = 7.38; p < .01). Duncan post-hoc testing showed that perceived workload for the low-complexity website (website 1) was significantly lower than the other two.

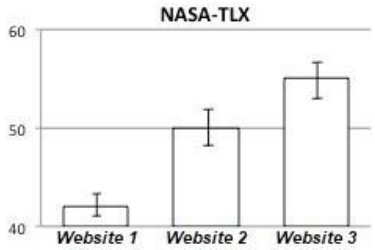

Figure 6. NASA-TLX scores per website.

The Nearest Neighbour Index was significantly different between websites ($F_{2,36}$ = 6.41; $p < .01$). Duncan post-hoc testing showed that the fixation pattern of the medium- and high-complexity websites (websites 2 and 3) were significantly more clustered than the low-complexity website.

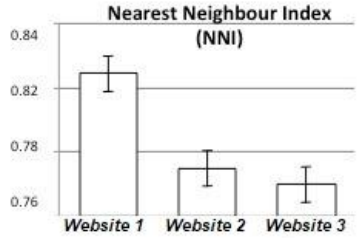

Figure 7. Nearest Neighbour Index per website.

Discussion and conclusion

Ensuring high usability of e-Governments tools, such as websites and web-based platforms plays an essential role in improving the efficiency in terms of costs and quality of services offered to citizens and businesses. Therefore, the evaluation of user performance and perceptions during the interaction with institutional websites has a key role both in the design phase of the interfaces and in their usability assessments. However, despite the growing interest of Public Administrations towards those topics, the relationship between complexity of information architecture, usability perceptions and mental workload imposed on the user has not been sufficiently investigated.

The objective of the present study was to evaluate the workload imposed on the user by "information abundant" websites and its relation with perceived usability. The idea that the complexity of information architecture can influence both the cognitive load imposed on user and his perceptions related to the pleasure of the user experience has been already suggested in the literature (e.g. Conklin, 1987). However, there is a lack of experimental research on this issue.

Here we have selected three existing Italian Government websites with different levels of information complexity to test the research hypothesis. Results indicated a consistency between satisfaction, usability and mental workload measures. Specifically, the websites associated to lower levels of mental workload (assessed by

both objective and subjective measures) received more positive usability evaluations were associated with a greater successes rate in the assigned search tasks, and participants reported their willingness in spreading the word about their usefulness.

The analysis of eye movements showed statistically significantly higher NNI values for the website 1, while lower values were reported for the websites 2 and 3. Considering the visual exploration strategies, this result highlights a less clustered fixations pattern for the website 1 and, on the contrary, a more clustered fixations pattern for the websites 2 and 3. Based on previous research (Camilli, Terenzi, & Di Nocera, 2008), when a task imposes high visual-spatial demand on the user -as in the case of an information search task- fixations clustering (i.e.: smaller NNI values) corresponds to greater mental workload experienced by the user. Subjects involved in this study have indeed experienced a greater mental workload while browsing the higher-complexity websites. This result is also supported by the subjective evaluations expressed using the NASA-TLX. Consistent results can be found in the usability evaluations expressed using the Us.E 2.0. scales, the "Handling" dimension (which is related to the information architecture) being the most problematic. Usability evaluations of a website are generally negative when users take too long to complete the task, make mistakes or fail in its execution (Nielsen & Levi, 1994; Nielsen, 1999; Palmer, 2002). In this study success rate was higher and completion time shorter for low-complexity website than the other two.

In conclusion, results obtained in this study underline the necessity and importance of integrating an estimate of users' mental workload during the design and evaluation of usability of complex websites such as those of Public Administrations. Some limitations that may have influenced the findings of this study should also be reported. A first limitation can be attributed to the scarce heterogeneity of the experimental sample in terms of age range and occupation. In fact, the age of the participants was between 46 and 65. Moreover, all the participants were Government employees who may have, in some way, benefited from their knowledge regarding the PA's websites structure and services. Although the type of processes investigated here can be considered common to all individuals, it could be useful to involve different types of users (e.g. experts vs. naives, typical vs. occasional users, etc.). Future studies could also include other measures (subjective and objective) in order to better understand the relation between the observed variables increasing the validity of results. In addition, the relationship between usability and mental workload could be influenced by both the information architecture and the layout of websites. A further objective could be to understand the specific role of these two variables. For example, conducting further studies contrasting the same information architecture with different layouts (e.g. the mobile version of the same websites), could allow greater generalization of the observed relationship between information architecture, mental workload and usability perceptions.

References

Bellenkes, A.H., Wickens, C.D., & Kramer, A.F. (1997). Visual scanning and pilot expertise: the role of attentional flexibility and mental model development. *Aviation, Space, and Environmental Medicine*, *68*, 569-579.

Brookings, J.B., Wilson, G.F., & Swain, C.R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology*, *42*, 361-377.

Bunecke, J.L. (1987). Quantifying some information processing requirements of the pilot's instrument crosscheck. In *Proceedings of the Human Factors Society Annual Meeting, Volume* 31 (pp. 1301-1305). Los Angeles: SAGE Publications.

Camilli, M., Terenzi, M., & Di Nocera, F. (2008). Effects of temporal and spatial demands on the distribution of eye fixations. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Volume 52* (pp. 1248-1251). Los Angeles: Sage Publications.

Conklin, J. (1987). Hypertext: an introduction and survey. *Computer*, *20*, 17 – 41.

Cowen, L., Ball, L.J., & Delin, J. (2002). An eye movement analysis of web page usability. In *People and Computers XVI-Memorable Yet Invisible* (pp. 317-335). London: Springer.

Di Nocera, F. (2013). *Usability Evaluation 2.0: Una descrizione (s)oggettiva dell'usabilità*. Roma: Ergoproject.

Di Nocera, F., Camilli, M., & Terenzi, M. (2007). A random glance at the flight deck: pilot's scanning strategies and the real-time assessment of mental workload. *Journal of Cognitive Engineering and Decision Making*, *1*, 271-285.

Di Nocera, F., Ranvaud, R., & Pasquali, V. (2015). Spatial pattern of eye fixations and evidence of ultradian rhythms in aircraft pilots. *Aerospace Medicine and Human Performance*, *86*, 647-651.

Ehmke, C., & Wilson, S. (2007). Identifying web usability problems from eye-tracking data. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI... but not as we know it-Volume 1* (pp. 119-128). British Computer Society.

Ephrath, A.R., Tole, J.R., Stephens, A.T., & Young, L.R. (1980). Instrument Scan—Is it an Indicator of the Pilot's Workload? In *Proceedings of the Human Factors Society Annual Meeting, Volume 24* (pp. 257-258). Los Angeles: Sage Publications.

Fedele, G., Fedriga, M., Zanuso, S., Mastrangelo, S., & Di Nocera, F. (2017). Can User Experience affect buying intention? A case study on the evaluation of exercise equipment. In: D. de Waard, A. Toffetti, R. Wiczorek, A. Sonderegger, S. Röttger, P. Bouchner, T. Franke, S. Fairclough, M. Noordzij, and K. Brookhuis (Eds.). *Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2016 Annual Conference*. (pp. 145-152). ISSN 2333-4959 (online). Available from http://hfes-europe.org

Fitts, P.M., Jones, R.E. & Milton, J.L. (1950), Eye Movements of Aircraft Pilots during Instrument-landing Approaches. *Aeronautical Engineering Review 9*, 24–29.

Garrett, J.J. (2010). *Elements of user experience, the: user-centered design for the web and beyond*. Pearson Education.

Goldberg, J.H., & Kotval, X.P. (1999). Computer interface evaluation using eye movements: methods and constructs. International Journal of Industrial Ergonomics, *24*, 631-645.

Goldberg, J.H., Stimson, M.J., Lewenstein, M., Scott, N., & Wichansky, A.M. (2002). Eye tracking in web search tasks: design implications. In *Proceedings of the 2002 symposium on Eye tracking research & applications* (pp. 51-58). ACM.

Goldberg, J.H., & Wichansky, A.M. (2003). Eye tracking in usability evaluation: A practitioner's guide. *In The Mind's Eye* (pp. 493-516). North-Holland.

Gullikson, S., Blades, R., Bragdon, M., McKibbon, S., Sparling, M., & Toms, E.G. (1999). The impact of information architecture on academic web site usability. *The Electronic Library*, *17*, 293-304.

Habuchi, Y., Kitajima, M., & Takeuchi, H. (2008). Comparison of eye movements in searching for easy-to-find and hard-to-find information in a hierarchically organized information structure. In *Proceedings of the 2008 symposium on Eye tracking research & applications* (pp. 131-134). ACM.

Hankins, T.C. & Wilson, G.F. (1998). A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight. *Aviation, Space & Environmental Medicine*, *69*, 360-367.

Hart, S.G. (2006). *NASA-task load index (NASA-TLX); 20 years later.* In *Proceedings of the human factors and ergonomics society annual meeting, Volume 50* (pp. 904-908). Los Angeles: Sage Publications.

Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology*, *52*, 139-183.

ISO 9241-11 (1998). Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)–Part II Guidance on Usability.

ISO 9241-210 (2009). Ergonomics of human system interaction-Part 210: Human-centred design for interactive systems.

Jacob, R.J.K. & Karn, K.S. (2003). Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises. In J. Hyona, R. Radach, and H. Deubel (Eds.), *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* (pp. 573–605). Amsterdam: Elsevier Science.

Kokini, C.M., Lee, S., Koubek, R.J., & Moon, S.K. (2012). Considering context: The role of mental workload and operator control in users' perceptions of usability. *International Journal of Human-Computer Interaction*, *28*, 543-559.

Kotval, X.P., & Goldberg, J.H. (1998). Eye movements and interface component grouping: An evaluation method. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Volume 45* (pp. 486-490). Los Angeles: SAGE Publications.

Longo, L. & Dondio, P. (2015). On the relationship between perception of usability and subjective mental workload of web interfaces. In *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2015 IEEE/WIC/ACM International Conference Volume 1 (*pp. 345-352). IEEE.

Majaranta, P., & Bulling, A. (2014). Eye tracking and eye-based human–computer interaction. In *Advances in physiological computing* (pp. 39-65). Springer: London.

May, J.G., Kennedy, R.S., Williams, M.C., Dunlap, W. P., & Brannan, J. R. (1990). Eye movement indices of mental workload. *Acta psychologica*, *75*, 75-89.

McGillis, L., & Toms, E.G. (2001). Usability of the academic library web site: implications for design. *College & research libraries*, *62*, 355-367.

Miller, J.M. (1973). Visual behavior changes of student pilots flying instrument approaches. In *Proceedings of the Human Factors Society 17th Annual Meeting* (pp. 208-214). Santa Monica (CA): Human Factors Society.

Miller, C.S., & Remington, R.W. (2004). Modeling information navigation: Implications for information architecture. *Human-computer interaction*, *19*, 225-271.

Morville, P., & Rosenfeld, L. (2006). Information Architecture 3.0. *Semantic studios*, *29*.

Nielsen, J. (1994). Usability inspection methods. In *Conference companion on Human factors in computing systems* (pp. 413-414). ACM.

Nielsen, J. (1999). *Designing web usability: The practice of simplicity.* New Riders Publishing.

Palmer, J.W. (2002). Web site usability, design, and performance metrics. *Information systems research*, *13*, 151-167.

Pan, B., Hembrooke, H.A., Gay, G.K., Granka, L.A., Feusner, M.K., & Newman, J.K. (2004). The determinants of web page viewing behavior: an eye-tracking study. In *Proceedings of the 2004 symposium on Eye tracking research & applications* (pp. 147-154). ACM.

Poole, A., Ball, L.J., & Phillips, P. (2005). In search of salience: A response-time and eye-movement analysis of bookmark recognition. In *People and computers XVIII—Design for life* (pp. 363-378). London: Springer.

Poole, A., & Ball, L.J. (2006). Eye tracking in HCI and usability research. *Encyclopedia of human computer interaction*, *1*, 211-219.

Reichheld, F.F. (2003). The one number you need to grow. *Harvard business review*, *81*, 46-55.

Reichheld, F.F., & Covey, S.R. (2006). *The ultimate question: Driving good profits and true growth.* Boston: Harvard Business School Press.

Sauro, J., & Lewis, J. R. (2010). Average task times in usability tests: what to report?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2347-2350). ACM.

Toms, E.G. (2002). Information interaction: Providing a framework for information architecture. *Journal of the Association for Information Science and Technology*, *53*, 855-862.

Van Orden, K.F., Limbert, W., Makeig, S., & Jung, T.P. (2001). Eye activity correlates of workload during a visuospatial memory task. *Human factors*, *43*, 111-121.

Wang, J., Antonenko, P., Celepkolu, M., Jimenez, Y., Fieldman, E., & Fieldman, A. (2018). Exploring Relationships Between Eye Tracking and Traditional Usability Testing Data. *International Journal of Human–Computer Interaction*, 1-12.

Wickens, C.D., & Hollands, J.G. (2000). *Engineering Psychology and Human Performance (3rd Edition)*. Upper Saddle River: Prentice Hall.