

Using Eye Tracking to Measure Overall Usability of Online Grocery Shopping Websites

Ali Bonyadi Naeini, Iran University of Science and Technology, Iran*

 <https://orcid.org/0000-0003-3119-551X>

Ali Golbazi Mahdipour, Iran University of Science and Technology, Iran

Rasam Dorri, Iran University of Science and Technology, Iran

ABSTRACT

This article examines usability evaluation methodologies, then presents a non-conscious behavioral indicator based on user eye movements and pupil dilation. The authors test how gender and online buying history affect the behavioral index's usability scores. This study uses three Iranian online food retailers. Thirty participants were asked to add things from predetermined grocery stores to virtual shopping carts before the experiment took them to the other two websites in a random order to collect eye movement data. Each group's presentation order was randomized. The number of fixations, number of saccades, total duration of fixations, scan-path length, pupil size, and task time were inversely linked with self-report usability measures. This research evaluates groups with different levels of online shopping expertise and gender based on experienced usability. Differences between groups suggest that user demographics affect usability.

KEYWORDS

Consumer Neuroscience, E-Commerce, Eye-Tracking, Neuromarketing, Online Shopping, Usability

1. INTRODUCTION

Shopping has become increasingly important as the internet has affected many other parts of people's daily lives all across the world. Electronic commerce websites like Amazon and eBay now do far more and on a much larger scale than traditional retail malls (Althafairi, Alhoumaida, Saxena, & Almsri, 2019; Cane & Parra, 2020). However, some circumstances, most notably the COVID-19 outbreak, have forced customers to use online buying sites even more frequently. (Villa & Monzón, 2021). B2C e-commerce enterprises have thrived in Iran and many other nations as a result (Malehmir, Maeen, & Jahangir, 2017). For example, according to Alexa.com, an Iranian e-commerce website called Digikala was the third most visited in Iran and the 140th most visited worldwide at the time this article was written (Analytics, 2019). Digikala.com's rapid development suggests that e-commerce is gaining traction in Iran. E-commerce sites are like massive shopping malls, attracting tens of thousands and

DOI: 10.4018/IJMCMC.326129

*Corresponding Author

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

in some cases, millions of customers each day. As a result, even the smallest changes to this massive store's layout can have a significant impact on sales.

Having said that, user-centered design is essential for e-commerce websites in the cut-throat business environment of today (Khosla, Damiani, & Grosky, 2003; Kramer, Noronha, & Vergo, 2000; Paknejad, Mosaddad, & Sadeghi Naeini, 2021; Sadeghi Naeini, Dalal, Mosaddad, & Karuppiah, 2018). To make sure they work effectively and efficiently, e-commerce websites should be regularly evaluated from different points of view. from their search engine optimization metrics (Hasan, Morris, & Proberts, 2009) to user interface design and usability (Sivaji, Downe, Mazlan, Soo, & Abdullah, 2011). One of the most important aspects of an e-commerce website is its usability (Sivaji et al., 2011), which has been of great concern to many researchers (Díaz, Rusu, & Collazos, 2017; Goh et al., 2013; Hasan et al., 2009; Singh, Malik, & Sarkar, 2016; Sivaji et al., 2011).

There are numerous explanations about what the term "usability" means (Bevan, 1995, 2008; Han, Yun, Kim, & Kwahk, 2000; Jokela, Iivari, Tornberg, & Electro, 2004; Jurek Kirakowski & Cierlik, 1998; Kwahk & Han, 2002; McNamara & Kirakowski, 2006; Thoma & Dodd, 2019). Nevertheless, the one suggested by ISO is presumably of more importance and has been recognized more widely. ISO 9241-11 (Jokela et al., 2004) defines usability 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." Han et al. in another established definition, delineate usability of consumer electronic products as "satisfying the users in terms of both the performance and the image and impression felt by them" (Han, Yun, Kwahk, & Hong, 2001).

According to the model proposed by ISO-9241, the usability of a system is comprised of three components:

- 1- *Effectiveness*, which is measured through the users' success rate in undertaking specified tasks using the system. In other words, the smaller the user mistake rate, the greater the effectiveness of the system under consideration.
- 2- *Efficiency* relates to the effort that a user should exert in order to complete a particular task using the system. This usability component is frequently measured using metrics such as Reaction Time, Task Completion Time, and Workload.
- 3- *Satisfaction*, Self-report assessments are often used to assess subjective and attitudinal components.

Besides these major components, usability models proposed by others suggest learnability (Constantine & Lockwood, 1999; Nielsen, 1994a; Preece et al., 1994; Scneiderman, 1992; Shackel, 2009) and memorability (Constantine & Lockwood, 1999; Nielsen, 1994a; Scneiderman, 1992; Shackel, 2009), safety (Roman, Ancker, Johnson, & Senathirajah, 2017), truthfulness (Ahuja, 2000; Atif, 2002), accessibility (Caldwell, Cooper, Reid, & Vanderheiden, 2008), universality (Seffah, Donyaee, Kline, & Padda, 2006), and usefulness (Seffah et al., 2006) attributes of the systems should be taken in to account when evaluating their usability. Learnability is measured by the average time a target user needs to learn using a system (Scneiderman, 1992). Yet memorability is measured through time that they can retain the knowledge or skills they need for using the system (Shackel, 2009). Needless to say, a system may be easy to learn but arduous to memorize or vice versa.

The safety of a system is concerned with its security, fault tolerance, environmental friendliness, and curbing any kind of harm to users and their resources (Seffah et al., 2006). The term "*Truthfulness*" refers to the expectation that if the system is given the correct input, it will create and deliver accurate outcomes (Ahuja, 2000; Atif, 2002; Friedman, Khan Jr, & Howe, 2000). But it still is a valid factor in many other systems as well. *Accessibility* deals with the capability of the system to be used by users of any kind of disability (Billi et al., 2010; Petrie & Kheir, 2007). *Universality*, relates to the extent to which the system can be used by users of different languages and cultural backgrounds (Clemmenssen, 2009; Cui, Wang, Pan, & Ni, 2020). And finally, *Usefulness* is concerned with addressing users' real needs in an appropriate way (Ishaq, Zin, Rosdi, Abid, & Ali, 2020; Snyder et al., 2021).

By and large, every system—consisting of interrelated elements forming a unified whole for a mutual purpose—can be a subject of usability evaluation. Yet, since the primary concern of this research is evaluating the usability of electronic commerce systems, hereafter, we will focus on the usability of online systems.

Because usability is largely experiential, relying solely on explicit self-report metrics runs the risk of misinterpretation because experiences might be conscious or unconscious (Likierman, 2006). That said, one of the methods which have been frequently used for measuring non-conscious experience is eye-tracking (Beattie & McGuire, 2012; Veto, Thomas, Alexander, Wemyss, & Mollon, 2020). Thanks to the recent development of eye-tracking technology, modern eye-tracking systems can now accurately and unobtrusively measure eye movements and pupil dilation of the users while they are working with the systems which their usability is to be measured. Thus, measuring the user experience through their eye movements is remarkably easier than it used to be. Here we investigate the association between well-established self report usability scales with eye-tracking metrics.

H1- There is a significant correlation between eye-movement metrics and standard self-report usability scales.

Furthermore, nowadays, the usability of systems is mostly regarded as an aspect of the user experience (Tomitsch, Janssen, Curwood, & Thomson, 2020) rather than a concrete attribute of the system, which is independent of the users' differences. However, a highly usable system for one group of users may not be as usable for another. Therefore, it is more crucial to create and apply usability indices that can effectively assess and contrast the experiences of various demographic groups. In this regard, eye-tracking has been argued to be a promising method.

Tupikovskaja-Omovie and Tyler (2021) compared the online shopping behavior of experienced and inexperienced consumers using eye-tracking. They compared how long experienced and inexperienced users spent at various points in the online shopping process. According to this study, unskilled users spend noticeably more time on the relevant sites during the majority of the stages.

Hence, we examine the effect of users' prior online-shopping experience on their usability of e-commerce websites by eye-tracking metrics.

H2- Users' prior online-shopping experience significantly affect their experienced usability of e-commerce websites as calculated by eye-tracking metrics.

In another study, Tupikovskaja-Omovie and Tyler (2020) investigated the user experience difference between men and women while undertaking online shopping tasks on a mobile device. In order to purchase two products within a predetermined budget, they collected eye-tracking data from 14 participants as they browsed the website of an online fashion company. The results from this research suggests that by and large, men tend to have fewer and longer fixations compared to women while shopping online. Therefore, we decided to measure the effect of user's gender usability while shopping on e-commerce websites.

H3- user's gender significantly affect their experienced usability of e-commerce websites as calculated by eye-tracking metrics.

2. BACKGROUND

2.1 Taxonomy of Usability Evaluation Methods

Since the usability assessment is needed for various tools and media, quite a lot of measures and scales have been developed for the evaluation of usability (Hartson, Andre, & Williges, 2001; Hasibuan,

Santoso, Yunita, & Rahmah, 2020; Hornbæk, 2010; Madan & Dubey, 2012). Plenty of the measures proposed and used for usability evaluation are subjective self-report scales (Friedrichs, Borojeni, Heuten, Lüdtke, & Boll, 2016; Wang, Terken, & Hu, 2014); others are behavioral metrics (Arinalhaq & Widyanti, 2019; Schmidt, Wittmann, & Wolff, 2019). That said, a review of the published usability studies by Forster, Hergeth, Naujoks, and Krems (2018) shows up until 2018 preponderance of research articles merely used self-report measures for studying usability of different systems. However, they revealed a trend of using behavioral measures of usability had already started by then.

In one of the proposed categorizations of usability evaluation methods has been put forth by (Bowman, Gabbard, & Hix, 2002), in this taxonomy, different methods are classified from three perspectives: (1) User involvement: methods that do not require users compared to those require user involvement for the evaluation process. (2) Context of evaluation: Generic versus application-specific methods (3) Type of results: methods providing qualitative results compared to those yielding quantitative results.

Additionally, different taxonomies have been proposed for general (Fitzpatrick, 1998; Ivory & Hearst, 2001) and specific uses (Asaddok & Ghazali, 2017; Rauf, Troubitsyna, & Porres, 2019; Rodrigues, de Borba Campos, & Zorzo, 2017). However, due to the broad look that (Bowman et al., 2002) provides, it has already drawn much attention and by the time has been cited 549 times according to Google Scholar.

Eye tracking has been used in previous studies for evaluating the usability of some types of online systems such as educational websites and E-learning (Zardari et al., 2021), online banking (Monica et al., 2019), and mobile health application (Cho et al., 2019). Furthermore, some researches (Modi & Singh, 2022) have used eye tracking for studying the usability of e-commerce websites. However, regardless of their importance, the usability of online grocery store has not been subject to appropriate investigation using eye tracking technique. In addition, the association between established self-report usability scales and eye-tracking metrics have not been investigated.

Also, To the best of our knowledge, no taxonomies have included electrophysiological and behavioral methodologies despite their expansion and advancement. We claim that these flaws render inapplicable Bowman et al. (2002) and other existing usability evaluation taxonomies. In order to address these imperfections, we propose the categorization rendered in Table 1.

For a broader understanding of each category of methods, see Table 2.

Newly discovered non-conscious procedures, regardless of conventional self-report metrics of usability, have not yet been thoroughly explored. Moreover, more thorough research on these metrics is required due to the potential cultural and contextual dependencies of these methodologies.

Having said that, the main goal of this research is to suggest an accurate, trustworthy, quick, and simple to use index for evaluating the usability of e-commerce websites. According to this study, the proposed index may reduce the need for large sample sizes in quantitative usability evaluation methods in order to obtain statistically significant results. Actually, this study evaluates the usability of the online grocery stores using eye-tracking method. Thus, this work belongs to the topic of neuromarketing, which has recently attracted a lot of interest from academics in several disciplines (Hsu & Chen, 2019; Juarez, Tur-Viñes, & Mengual, 2020; Nilashi et al., 2020; Rawnaque et al., 2020).

The Website Analysis and Measurement Inventory (WAMMI), a 20-item and five-point Likert scale inventory (Forçan, Abe, de Lima, & Nascimento, 2020; J Kirakowski, 1998), and the System Usability Scale (SUS), a 10-item and five-point Likert scale questionnaire (Brooke, 1996b; Holden, 2020; Mol et al., 2020) are the two major self-report measures used in this study. We excluded one of the WAMMI inventory items because it was unrelated to functions of e-commerce websites. With 19 items total, the WAMMI scale was utilized in this study. Novel taxonomy for usability evaluation methods, and summary of usability evaluation methods in Table 1 and Table 2 respectively.

Table 1. A novel taxonomy for usability evaluation methods

User Involvement			Consciousness			
			Conscious	Non-Conscious		
	Requires users	Performance	<ul style="list-style-type: none"> • Bare eye behavior observation (time of successful task completion, time on task, mouse counting, error tracking, behavior coding) 	<ul style="list-style-type: none"> • Electrophysiological methods (GSR, EEG, EMG, HRV) • Computer assisted behavior observation (eye-tracking, facial expression analysis, and mouse tracking) 	quantitative	Type of research
			<ul style="list-style-type: none"> • Bare eye behavior observation (e.g. behavior coding) 	<ul style="list-style-type: none"> • Eye tracking visualization (e.g. heat-maps) • Mouse tracking visualizations (e.g. heat-maps) 	qualitative	
		Self-report	<ul style="list-style-type: none"> • Survey scales • Standardized scale questionnaires • MADM 	<ul style="list-style-type: none"> • Implicit association measures 	quantitative	
			<ul style="list-style-type: none"> • Open-ended questionnaires • Interview • Focus group • Interview (structured, semi-structured, and unstructured) • Focus groups • Think Aloud Protocols (CTA, AI, and SC) • Information architecture (Tree testing, card sorting, click testing, etc) 	-	qualitative	
	Does not require users	Model / expert opinion	<ul style="list-style-type: none"> • Generic performance model methods (such as Fitt's law) • Cognitive walkthrough • Application-specific performance model methods (such as GOMS) • Heuristic Evaluation methods 	-	qualitative	

2.2 Eye-Tracking in Usability Evaluation

The state-of-the-art eye-tracking systems accurately and precisely measure the eye movements of the participants, their pupil dilation, and their distance to the object they are looking at and the position of their head in the three-dimensional space (Klaib, Alsrehin, Melhem, Bashtawi, & Magableh, 2021; Krebs et al., 2021; Niu, Zhou, & Bai, 2021). This information provides highly valid insights into cognitive processes as well as the genuine experience of the users. Although many researchers from various fields have used eye-tracking as a behavioral measure, here we merely focus on the most relevant works pertaining to the use of eye-tracking for the usability evaluation of online shopping platforms.

Table 2. Summarization of usability evaluation methods

Category	Method Type	Method	Definition	Ref.	
Self-report	Open-ended questionnaire	Open-ended questionnaire	Open-ended questions are usually used alongside other self-report measures in order to gain deeper insight into the system.	(Albert & Tullis, 2013; Wilson, 2013)	
	Multiple-choice questionnaires	Multiple-choice questionnaires	These measures render some predefined options to the participants and require them to choose one or more options.	(Albert & Tullis, 2013)	
	Standardized scales		System Usability Scale (SUS)	Scale measures, after each of the standardized questions, present the user with a range of numbers in which the two ends of the range are antitheses of each other. Participants can choose the respective number that most accurately represents their answer to the question. There have been quite a few standard scales with a different numbers of items developed and suggested for different types of systems.	(Brooke, 1996a; Mol et al., 2020; Sauro & Lewis, 2016; Vlachogianni & Tselios, 2021)
			Usability Magnitude Estimation (UME)		(Assila & Ezzedine, 2016; Rich & McGee, 2004; Sauro & Dumas, 2009)
			Website Analysis and Measurement Inventory (WAMI)		(Ahmad, Hussain, Flayyih, Abdulwahab, & Sabri, 2017; Sauro, 2015; Tullis & Stetson, 2004)
			Questionnaire for User Interaction Satisfaction (QUIS)		(Karoulis, Sylaiou, & White, 2006; Rotaru, Vert, Vasiu, & Andone, 2020; Tullis & Stetson, 2004)
			After Scenario Questionnaire (ASQ)		(Lattie et al., 2020; J. R. Lewis, 1991)
			Computer System Usability Questionnaire (CSUQ)		(Al-Hassan, AlGhannam, Naser, & Alabdulrazzaq, 2021; Alhadreti, 2021)
	Multiple Attribute Decision Making (MADM)		Analytic Hierarchy Process (AHP)	Different MADM methods have been used for spotting the most important issues among a range of identified issues or the most critical areas of the system that users regard as more important. In these methods, users fill out some questionnaires, and weights of the attributes and/or their priorities are calculated through the answers.	(Adepoju, Oyefolahan, Abdullahi, & Mohammed, 2020)
			Analytic Network Process (ANP)		(Al-Zahrani, 2020)
			Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS)		(Kumar et al., 2020)
	Think Aloud Protocols		Classical Think Aloud (CTA)	In Think Aloud Protocols, users are asked to undertake specific tasks using a system and meanwhile speak out loud and elaborate on their thoughts regarding various aspects of the system.	(Alhadreti & Mayhew, 2017; Elling, Lentz, & de Jong, 2011; Ji & Rau, 2019)
			Active Intervention (AI)		(Alhadreti & Mayhew, 2017)
			speech communication (SC)		(Alhadreti, 2016; Alhadreti & Mayhew, 2017; Fan, 2019)
	Implicit self-report		Implicit Association Test (IAT)	Implicit association tests measure how closely are different concepts and items associated with each other by users. The associations in this method are calculated through reaction times.	(Actis-Grosso, Capellini, Ghedin, & Tassistro, 2021)
	Information architecture		Tree testing	These methods aim at matching the structure of the information presented in the system to users' expectations. In these methods, users are asked to sort out the items of the system in a way that makes the most sense to them.	(Friebert, 2017; Soranzo & Cooksey, 2015; Sripathi & Sandru, 2013)
			Card sorting		(K. M. Lewis & Hepburn, 2010; Zimmerman & Akerele, 2002)
			Click testing		(Hassan, 2020)

continued on following page

Table 2. Continued

Category	Method Type	Method	Definition	Ref.
Behavioral	Computer assisted observation	Eye-tracking	These methods use a range of hardware and software packages for investigating users' behaviors, usually with a sampling frequency of greater than 15 Hz.	(Ehmke & Wilson, 2007; Goldberg & Wichansky, 2003; Nielsen & Pernice, 2010)
		Facial expression		(Chao, 2021; Märtin, Bissinger, & Asta, 2021; Savela-Huovinen, Toom, Knaapila, & Muukkonen, 2021; Veilleux et al., 2020)
		Pressure-sensitive mouse		(Dennerlein, Becker, Johnson, Reynolds, & Picard, 2003; Reynolds, 2005)
		Mouse tracking		(Arroyo, Selker, & Wei, 2006)
	Bare eye Observation	Mouse click counting	Bare eye observation methods are more conventional and do not necessarily require any specific electronic equipment; however, in some cases, a computer might be used to further facilitate the research.	(Tomeo, 2012)
		Error counting		(Page, 2011)
		Time-on-task		(Karweit, 1988)
		Time of successful task completion		(Butnampetch, Sasithonwan, Teeranan, & Chintakovid, 2020; Horton, 2011)
		Behavior coding		(Cook & Meyer, 2017; Geisen & Murphy, 2020)
	electrophysiological	Central Nervous System	Electroencephalography (EEG)	Electrophysiological methods have recently been used for usability evaluations through users' non-conscious experiences. These methods either work through measuring the electrical activity of the brain (EEG) or indirect peripheral organs.
Peripheral Nervous System		Galvanic Skin Response (GSR)	(do Amaral, Ferreira, Aquino, & de Castro, 2013; Falkowska, Sobiecki, & Pietrzak, 2016; Wenzel et al., 2015)	
		HeartRate Variations (HRV)	(Trimmel, Meixner-Pendleton, & Haring, 2003; Ward & Marsden, 2003)	
		Electromyography (EMG)	(Benedek & Hazlett, 2005)	
Experts' opinion	Heuristic	Jakob Nielsen's Usability Heuristics	Heuristic usability evaluation methods require experts to rate different critical attributions of the current state of the system. Different heuristic models have different criteria and attributions.	(Nielsen, 2011; Nielsen & Molich, 1990)
		Arnie Lund's Heuristics		(Lund, 1997)
		Bruce Tognazzini's Heuristics		(Tognazzini, 2014)
		Ben Shneiderman's Heuristics		(Shneiderman, 1986)
	Cognitive walkthrough	Sketch-based	In the walk through methods, usually, experts that are not yet familiarized with the system try to go through different tasks and extract the problems that an average user would face while carrying out the task. This method is very low cost but not highly accurate.	(C. Lewis & Wharton, 1997; Rieman, Franzke, & Redmiles, 1995)
		Prototype-based		(Liu, Osvalder, & Dahlman, 2005)
		Product-based		(Dewi, Dantes, & Indrawan, 2020)
		Pluralistic Walkthrough		(Nielsen, 1994b)

continued on following page

Table 2. Continued

Category	Method Type	Method	Definition	Ref.
Discussion-based	interview	Unstructured	In these methods, one or more users discuss different aspects of a system questioned by an interviewer or moderator. In some methods, there might be more than one moderator, or some of the users might be assigned to the role as moderators. In other methods two moderators with opposing perspectives ignite a discussion to elicit users' responses.	(Pittig, Hoyer, & Noack, 2021)
		Semi-structured		(Asghar, Cang, & Yu, 2018)
		structured		(Lagha et al., 2020)
	Focus group	Single Focus Group		(Caplan, 1990)
		Mini Focus Group		(McCloskey, Johnson, Bekelman, Martin, & Bellows, 2019)
		Two-Way Focus Group		(Chuka-Maduji, 2021)
		Dual Moderator Focus Group		(Goertzen & Bakkalbasi, 2016)
Dueling Moderator Focus Group	(Chuka-Maduji, 2021)			
Model methods	GOMS	CPM-GOM	Model-based methods use some specific criteria for predicting the difficulty of carrying out a task an action using the system. GOMS and Fitt's models are among the most famous models used both in academic and commercial research.	(Gong & Elkerton, 1990; D. Kieras, 1997; D. E. Kieras, 1999)
		NGOMSL		(D. Kieras, 1997)
		KLM-GOMS		(Rice & Lartigue, 2014)
	Fitt's law	Welford model		(Hoffmann, 2013)
		Mackenzie's variant		(Hoffmann, 2013)
		Conventional Fitt's law		(Hoffmann, 2013)

Eye-tracking is sometimes used in usability studies to analyze only a portion of an online commerce platform, with the goal of improving key and more frequently used elements of the website or mobile application. A/B testing (Siroker & Koomen, 2013) is the most common practice in this approach. For example (Das, McEwan, & Douglas, 2008) considered the alignment of field labels in user forms and compared the top, left, and right alignment of the labels regarding their effect on the usability of the user forms in an e-commerce website. According to their research, users who typically take longer to complete forms fill them out more quickly when labels are aligned to the right, but top alignment is quicker for users who typically take shorter amounts of time (Chen & Pu, 2010) concentrated on the design of the product recommenders regarding their interface structure. They compared three recommender interface designs with different structures. Per the study's findings, users are likely to focus on recommended products more and for a longer period of time if they are arranged in a quadrant layout. (Hwang & Lee, 2020).

From a novel perspective, Luan, Yao, Zhao, and Liu (2016) compared the user's visual behavior while investigating search and experience products. By search products, they meant items like laptops and mobile phones that the buyer would probably pay closer attention to attributes. And by "experience product," they meant those goods—like clothing and shoes—where users are most likely to consider the experiences of previous purchasers.

3. MATERIALS AND METHODS

This section will cover the demographic information and inclusion criteria used to select the subjects, the hardware and software used to collect eye-tracking data, the self-report scales used to gauge perceived usability of the tested websites, the protocol that participants had to follow, the algorithms used to preprocess the data, and the statistical tests that were performed on the preprocessed data.

3.1 Participants

Thirty right-handed participants (15 female) took part in this study (average age of 29, $SD=4.45$). 19 of the subjects were pursuing or had already completed their Master's degree, 8 were pursuing or had already completed their Ph.D., and the remaining 3 were pursuing or had already completed their bachelor's degree. Participants were recruited through convenience sampling with the previous online shop experiences. We made our sample size decision in accordance with earlier research that examined the dependability of eye-tracking data for various sample sizes (Zhang & Liu, 2017). In the study conducted by Zhang and Liu (2017), for instance, it was found that the eye-tracking data for static picture stimuli is fairly constant for as many as 12 subjects and exceeds the saturation levels at 16 subjects. Thus, we assume a sample size of 30 subjects will be sufficient for the current study. We made an effort to find individuals with both low and high levels of prior internet shopping experience. Their online shopping experience was measured using a five-point scale (1- In the past one year, I made no online purchases 2- In the past one year, I made one to three online purchases 3- In the past one year, I made four to eight online purchases 4- In past one year I made 9 to 12 online purchases 5- In past one year I made more than 12 online purchases). The average online shopping experience score of the participants was 2.76 ($SD= 1.77$, $Min= 0$, and $Max= 5$). According to the self-report website familiarity measured after the experiments, none of the participants were familiar with the websites or had previously used them. However, four subjects were familiar with the brand of website A, all of them were familiar with the brand of website B, and none of the subjects were familiar with the brand of website C (see apparatus and stimuli section for further information on websites). Participants reported that their familiarity with the brands was caused by either advertisements or familiarity with the parent brand (websites A and B were subsidiaries of two major brands). That said, regardless of their familiarity with the brand, participants had not been previously exposed to the websites.

3.2 Apparatus and Stimuli

The Tobii X2-30 eye tracker, with a sampling rate of 30 Hz, has been used in this study along with Tobii Studio 3.4.5 eye-tracking software package for stimulus presentation, data recording, and data preprocessing. For subjective recording evaluations of the participants regarding each of the e-commerce websites that were used as stimuli, we used two standard scales, namely SUS (10 items) and WAMMI (19 items). Although the original WAMMI questionnaire has 20 items, due to the communicational function of one of the items in the WAMMI inventory—asking about how well the users can communicate to those they wish using the website—we excluded it.

Websites of three Iranian online grocery stores—Okala, Snappmarket, and Tezolmarket—were used as the main test stimuli. From here on, we will name them as websites A, B, and C respectively. In order to minimize the effect of familiarity, the websites were chosen as though participants have minimum experience of visiting those websites. Thus, we did not include Digikala's online supermarket, which is the most popular Iranian online grocery seller.

3.3 Protocol

Subjects were questioned about their demographics, level of internet experience, history of e-commerce purchases, and familiarity with the tested websites. The demographic questionnaire also asked the participants if they were right-handed or left-handed and only right-handed participants who were not familiar with the target websites were invited to take part in the experiments. Upon entering the lab, where all of the experiments were conducted, subjects received a debriefing about the general procedure of the experiment as well as what types of data we would gather and how we would handle the data. Prior to data gathering, an electronic written informed consent was received from the participants according to the Declaration of Helsinki. Three tasks were established to simulate online grocery shopping. That is 1- purchasing dairy product 2- purchasing canned food, 3- purchasing drinks. Subjects were instructed to imagine themselves in a real typical situation where they need to buy these items. The assignment of the tasks to each of the websites, as well as the order in which

different participants were presented with the websites, were randomized. We used Tobii Studio's Sequence option for randomizing the order of stimuli and assignment of the tasks. Participants were asked to answer the SUS and modified WAMMI questions for the website they had just used based on their experiences after completing each task and before moving on to the next one.

3.4 Analysis

In the analysis stage of this study, by Tobii Studio software, the recorded eye-tracking data was segmented and unwanted parts of the recordings were excluded. The Tobii IV-T filter (Olsen, 2012), as implemented in Tobii Studio, was used for preprocessing the eye movements and classifying fixations. Afterward, areas of interest (AOIs) were defined over the whole page for the segmented parts of the recording. Then, participants were divided into groups by their gender and online purchasing experience. The eye-tracking metrics including pupil size, fixation count, duration of fixations, total duration of fixations, saccade count, saccade amplitudes, and duration of segments, were then exported from the software. Later by summing up amplitudes of saccades and using durations of segments, we calculated scan path length and task times, respectively. We used the Spearman correlation test to determine the relationship between behavioral and self-report measures because the data had a non-normal distribution, as determined by the Shapiro-Wilk normality test.

4. RESULTS

The initial goal of this study was to look into the relationship between eye-tracking metrics and already established self-report usability scales. The Shapiro-Wilk normality test showed the data were not normally distributed. Thus, in this phase of the study, the Spearman correlation test was used. That is, the correlation between behavioral and eye-tracking parameters and the SUS and WAMMI was calculated. Eye-tracking and behavioral measures included in this statistical test were fixation count, duration of fixations, total duration of fixations, saccade count, pupil size, task time, and scan path length. Since the self-report measures were highly correlated to each other, and there were 29 single questions asked from subjects for every website, we chose to report the test results for averaged scale scores here. However, the complete test results are available in the appendix. Table 3 presents the findings of the correlation test on the average score.

Regarding the results reported in Table 3, fixation count, fixation duration, total fixation duration saccade count, pupil size, task time, and scan path length are negatively correlated to the Website Analysis and Measurement Inventory scale. The correlation coefficients for the System Usability Scale are almost the same as WAMMI but for fixation duration. According to the Spearman correlation coefficients, fixation count, saccade count, and scan path metrics have the strongest association with the self-report usability scores.

Having yielded statistically significant results in the correlation test between behavioral metrics and self-report measures, we then used the behavioral metrics as a ground for comparing the visual behavior of men and women as non-conscious usability measures. So, in the second phase of this study, we used a t-test to compare two groups of men and women. Results from this test are reported in Table 4.

The results shown in Table 4 show that there was a significant difference between men and women's visual behavior and, consequently, usability when making online grocery purchases. Having looked at the means of the eye-tracking and behavioral metrics, it's seen that women score higher fixation count, and saccade counts, longer duration of fixations, and larger pupil dilation. Descriptive statistics comparing the eye-tracking and behavioral metrics among men and women have been reported in Table 5.

When we consider the eye-tracking metrics from the earlier analysis phase, we can understand why women appear to have lower usability experience scores than men. This might be due to women's need for more details when making decisions in order to reduce their decision risk (Arnsten, 1998;

Table 3. Association of behavioral metrics with usability scales' score

Measure	Stat	FC	FD	TFD	SPL
WAMI	<i>Rho</i>	-.731**	-.189*	-.630**	-.737**
	<i>P</i>	< 0.001	0.037	< 0.001	< 0.001
	<i>N</i>	90	90	90	90
SUS	<i>Rho</i>	-.699**	-0.158	-.601**	-.677**
	<i>P</i>	< 0.001	0.068	< 0.001	< 0.001
	<i>N</i>	90	90	90	90
Measure	Stat	SC	PS	TT	-
WAMI	<i>Rho</i>	-.748**	-.404**	.451**	
	<i>P</i>	< 0.001	< 0.001	< 0.001	
	<i>N</i>	90	90	90	
SUS	<i>Rho</i>	-.695**	-.373**	.387**	
	<i>P</i>	< 0.001	< 0.001	< 0.001	
	<i>N</i>	90	90	90	

Table 4. Results of t-test on online-shopping experience

Metrics	<i>F</i>	<i>Sig</i>	<i>t</i>	<i>df</i>	<i>Sig</i>
FC	0.134	0.717	4.942	28	< 0.001
FD	0.016	0.899	3.942	28	< 0.001
TFD	0.000	0.984	5.079	28	< 0.001
SC	0.210	0.650	3.994	28	< 0.001
PS	0.089	0.767	4.511	28	< 0.001
TT	0.750	0.394	5.380	28	< 0.001
SPL	0.003	0.955	3.021	28	0.005

Keinan, 1987; Keinan, Friedland, & Ben-Porath, 1987; Porcelli & Delgado, 2009). Thus, we assume women try to get more information when they are observing the web pages so that they can avoid taking risks when deciding to buy a product.

Concerning the other hypothesis of the present study, we conducted a second t-test to examine the effect of prior online shopping experience on the participant's visual behavior while purchasing groceries. Results of this t-test have been reported in Table 6.

Results from this test suggests those participants with higher online shopping experience have fewer fixations and saccades, shorter fixation duration, total fixation duration, and scan path length, as well as lower pupil dilation when undertaking the online purchasing task. In other words, the results indicate that users with greater e-shopping experience perceive the website to be more usable than those with less e-shopping experience. Descriptive statistics comparing two groups of subjects by their online-shopping experience have been reported in Table 7.

According to the results reported in Table 7. The higher the participants' experience with e-commerce websites, the fewer the number of fixations and saccades, and the shorter the scan path.

Table 5. Descriptive statistics of the metrics across genders

Metric/Gender		N	Mean	SD	SE Mean
FC	F	15	510.67	32.27	8.33
	M	15	393.73	32.83	8.48
FD	F	15	0.43	0.02	0
	M	15	0.35	0.02	0
TDF	F	15	216.87	13.23	3.41
	M	15	135.74	10.76	2.78
SC	F	15	679.67	55.28	14.27
	M	15	542.6	54.03	13.95
PS	F	15	2.77	0.11	0.03
	M	15	2.26	0.08	0.02
TT	F	15	243.31	11.8	3.05
	M	15	188.75	12.31	3.18
SPL	F	15	1403.18	169.95	43.88
	M	15	1136.72	158.47	40.92

Table 6. T-test results comparing men's and women's visual behavior when shopping online

Metrics	F	Sig	t	df	Sig
FC	0.084	0.774	9.839	28	< 0.001
FD	0.285	0.597	12.886	28	< 0.001
TFD	0.304	0.586	18.428	28	< 0.001
SC	0.108	0.744	6.867	28	< 0.001
PS	2.296	0.141	14.286	28	< 0.001
TT	0.072	0.791	12.389	28	< 0.001
SPL	0.018	0.894	4.441	28	< 0.001

The visualized results of the eye-tracking experiment, namely Heat maps and Gaze Plots for the main pages of the three e-commerce websites tested in this study, are as seen in Figure 1.

Figure 1 illustrates the visual performance disparity between three online grocery stores. In order to yield more reliable results, we chose the three test websites regarding their usability evaluated by a user experience expert. We aimed to include websites with low, medium, and high usability performance. The patterns of eye movement depicted in Figure 1 indicate that we were largely successful. We reason if the website is more usable, the fixations should be somewhat evenly distributed with higher density on target products. As seen in Figure 1 for website B (lower usability), participants mostly attended to the top of the page where they expected a product classification tree and paid rather lower attention to the other parts of the user interface. For website A (medium usability) and website C (higher usability), participants attended to the target products more often and their visual attention was rather smoothly distributed.

Table 7. Descriptive statistics comparing participants with high/low online-shopping experiences

Metrics	Exp.*	N	Mean	SD	SE Mean
FC	L	14	500.64	50.06	13.38
	H	16	409.81	50.35	12.59
FD	L	14	0.41	0.04	0.01
	H	16	0.36	0.04	0.01
TFD	L	14	207.55	32.21	8.61
	H	16	148.96	30.91	7.73
SC	L	14	666.86	67.04	17.92
	H	16	562.38	75.11	18.78
PS	L	14	2.7	0.2	0.05
	H	16	2.35	0.23	0.06
TT	L	14	238.64	16.57	4.43
	H	16	196.25	25.05	6.26
SPL	L	14	1379.79	192.48	51.44
	H	16	1173.84	180.76	45.19

* Exp: experience

Figure 1. Visualized eye movement patterns on three e-commerce websites tested (from left to right and top to bottom, Website A, Website B, Website C gaze plots; Website A, Website B, and Website C heat maps)



5. DISCUSSION

This study shows there might be a novel usability metric available for the evaluation of usability in e-commerce websites. This metric, opposing with the self-report metrics, may not be easily infected by the participants' conscious or unconscious deviation from the truth. Thus, such a measure will be way more valid, and we argue that if the participants sampled to be tested chosen according to the real target population, the results indicated by this novel metric can be reliable as well. Regarding the previous research undertaken on the effect of website user interface design on the eye movement patterns of the visitors, (Goldberg & Wichansky, 2003) suggests the fixation count and saccades users have while working with a user interface are affected by the visual complexity of the website. We argue that the higher the visual complexity of a website, the lower the usability score of that website thus we assumed that fixation count and saccade count would be negatively correlated with the self-report usability scales. This fact was proven by the results of the current study. According to a review done by (Ehmke & Wilson, 2007), the scan path length also is negatively correlated with the overall usability of a website which has been proven with results from the current research as well. Regarding the effect of gender on eye movement, a prior research suggests young boys and girls code activities—as measured using eye-tracking differently, but they reasoned more empirical evidence is needed for substantiating such hypothesis (Papavlasopoulou, Sharma, & Giannakos, 2020). Another study showed women have more saccades and length scan path length compared to men when viewing media (Sargezeh, Tavakoli, & Daliri, 2019). Thus, the current research is in line with the literature. This study has some limitations, much like any experimental research, and we recommend that future studies make up for them. Nowadays, we cannot ignore the prevalence of mobile phones and how they facilitate our access to the internet and the e-commerce sector; consequently, we believe that examining the usability of mobile versions of websites is an important and necessary area of study. Due to time and financial constraints, we were unable to include mobile versions of the websites in our experiment and hence cannot discuss them in this study. The eye-tracking equipment employed in this investigation had a sampling rate of 30 Hz, which is not quite high for a fully accurate computation of saccades. Additionally, we picked the websites based on how subjectively usable they were. It is strongly advised that later studies conduct a pre-test on the usability metrics of the websites in order to ensure that the websites have both high and low usability metrics. Additionally, creating fictional websites to assess their usability is advised for the researchers. By doing this, companies may also better manage the consequences of brand and website familiarity.

This research failed to control the participants' familiarity with the brands, which may have introduced some errors into the findings. Despite the fact that it controlled for website familiarity (none of the participants had ever used either of the websites before), this research did not control for brand familiarity. On the other hand, we reasoned that since purchasing random food products in unfamiliar online grocery stores is more of a cognitive task than an affective task, the familiarity with the brand and attitude towards the brand should not play a significant role in the overall experienced usability of the websites. Yet we suggest future studies also control these brand-related variables. Also, Zhang and Liu (2017) found saturation with 16 individuals seeing static stimuli, however other studies suggest a larger sample size can increase reliability. Consequently, future studies should use larger samples.

6. CONCLUSION

Eye-tracking metrics such as fixation count, saccade count, fixation duration, total fixation duration, and scan path length, along with pupil size, can be effectively used as an alternative to self-report scales, which are prone to producing inaccurate results. These findings were discovered as a result of a study that was conducted. It shows these metrics are negatively correlated with the self-report usability measures. Since these measures are mostly unconscious, participants may not control

them, and thus the results drawn from these metrics are more valid compared to those resulting from self-report measures. Furthermore, to evaluate the usability of a website is a valid and reliable way, researchers even if they don't use eye-tracking metrics, usually have to have participants go through multiple tasks and then have them fill the self-report questionnaire (Davis, Gardner, & Schnall, 2020; Unrau & Kray, 2019). Thus, using eye-tracking metrics proposed by the current study instead can reduce the duration of the research. Results from the current study also suggest women have a harder time finding and choosing the items to be added to their cart compared to men. The findings of this study also demonstrate that the usability of an e-commerce website increases with past online buying experience.

COMPETING INTERESTS

The authors of this publication declare there are no competing interests.

FUNDING AGENCY

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Funding for this research was covered by the authors of the article.

REFERENCES

- Actis-Grosso, R., Capellini, R., Ghedin, F., & Tassistro, F. (2021). *Implicit Measures as a Useful Tool for Evaluating User Experience*. Paper presented at the International Conference on Human-Computer Interaction. doi:10.1007/978-3-030-78092-0_1
- Adepoju, S. A., Oyefolahan, I. O., Abdullahi, M. B., & Mohammed, A. A. (2020). *Multi-criteria decision-making based approaches in website quality and usability evaluation: A systematic review*. Research Gate.
- Ahmad, A., Hussain, A., Flayyih, O. H., Abdulwahab, W., & Sabri, M. I. (2017). Utilizing WAMMI components to evaluate the usability of E-commerce website. *Journal of Telecommunication, Electronic and Computer Engineering*, 9(2-11), 139-143.
- Ahuja, V. (2000). Building trust in electronic commerce. *IT Professional*, 2(3), 61–63. doi:10.1109/6294.846215
- Al-Hassan, A. A., AlGhannam, B., Naser, M. B., & Alabdulrazzaq, H. (2021). An Arabic Translation of the Computer System Usability Questionnaire (CSUQ) with Psychometric Evaluation Using Kuwait University Portal. *International Journal of Human-Computer Interaction*, 37(20), 1–8. doi:10.1080/10447318.2021.1926117
- Al-Zahrani, F. A. (2020). Evaluating the Usable-Security of Healthcare Software Through Unified Technique of Fuzzy Logic, ANP and TOPSIS. *IEEE Access : Practical Innovations, Open Solutions*, 8, 109905–109916. doi:10.1109/ACCESS.2020.3001996
- Albert, W., & Tullis, T. (2013). *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Newnes.
- Alhadreti, O. (2016). *Thinking About Thinking Aloud*. Science Open.
- Alhadreti, O. (2021). Assessing academics' perceptions of blackboard usability using SUS and CSUQ: A case study during the COVID-19 pandemic. *International Journal of Human-Computer Interaction*, 37(11), 1003–1015. doi:10.1080/10447318.2020.1861766
- Alhadreti, O., & Mayhew, P. (2017). To intervene or not to intervene: An investigation of three think-aloud protocols in usability testing. *Journal of Usability Studies*, 12(3), 111–132.
- Althafairi, B., Alhoumaida, N., Saxena, M., & Almsri, Z. (2019). Case study-AMAZON. *Journal of the Community Development in Asia*, 2(2).
- Analytics, A. W. (2019). Alexa. com. In.
- Arinalhaq, F., & Widyanti, A. (2019). The role of national culture on the design and usability of e-commerce websites: An Indonesian case. *Journal of Physics: Conference Series*. doi:10.1088/1742-6596/1175/1/012192
- Arnsten, A. F. (1998). The biology of being frazzled. *Science*, 280(5370), 1711–1712. doi:10.1126/science.280.5370.1711 PMID:9660710
- Arroyo, E., Selker, T., & Wei, W. (2006). Usability tool for analysis of web designs using mouse tracks. *CHI'06 extended abstracts on Human factors in computing systems*. doi:10.1145/1125451.1125557
- Asaddok, N., & Ghazali, M. (2017). Exploring the usability, security and privacy taxonomy for mobile health applications. In *2017 International Conference on Research and Innovation in Information Systems (ICRIIS)*. IEEE. doi:10.1109/ICRIIS.2017.8002472
- Asghar, I., Cang, S., & Yu, H. (2018). Usability evaluation of assistive technologies through qualitative research focusing on people with mild dementia. *Computers in Human Behavior*, 79, 192–201. doi:10.1016/j.chb.2017.08.034
- Assila, A., & Ezzedine, H. (2016). Standardized usability questionnaires: Features and quality focus. *Electronic Journal of Computer Science and Information Technology*, 6(1).
- Atif, Y. (2002). Building trust in e-commerce. *IEEE Internet Computing*, 6(1), 18–24. doi:10.1109/4236.978365
- Beattie, G., & McGuire, L. (2012). See no evil? Only implicit attitudes predict unconscious eye movements towards images of climate change. *Semiotica*, 2012(192), 315–339. doi:10.1515/sem-2012-0066

- Benedek, J., & Hazlett, R. (2005). *Incorporating facial emg emotion measures as feedback in the software design process*. Proc. Human Computer Interaction Consortium.
- Bevan, N. (1995). Usability is quality of use. In *Advances in Human Factors/Ergonomics* (Vol. 20, pp. 349–354). Elsevier.
- Bevan, N. (2008). Classifying and selecting UX and usability measures. In *International Workshop on Meaningful Measures: Valid Useful User Experience Measurement*. Research Gate.
- Billi, M., Burzagli, L., Catarci, T., Santucci, G., Bertini, E., Gabbanini, F., & Palchetti, E. (2010). A unified methodology for the evaluation of accessibility and usability of mobile applications. *Universal Access in the Information Society*, 9(4), 337–356. doi:10.1007/s10209-009-0180-1
- Bowman, D. A., Gabbard, J. L., & Hix, D. (2002). A survey of usability evaluation in virtual environments: Classification and comparison of methods. *Presence (Cambridge, Mass.)*, 11(4), 404–424. doi:10.1162/105474602760204309
- Brooke, J. (1996a). SUS-A quick and dirty usability scale. *Usability Evaluation in Industry*, 189(194), 4–7.
- Butnampetch, P., Sasithonwan, P., Teeranan, B., & Chintakovid, T. (2020). Usability Studies of E-Commerce Checkout Process: A Perspective from Thailand. In *International Conference on Human-Computer Interaction*. Springer. doi:10.1007/978-3-030-50341-3_24
- Caldwell, B., Cooper, M., Reid, L. G., & Vanderheiden, G. (2008). Web content accessibility guidelines 2.0. *W3C Recommendation*, 11.
- Cane, M., & Parra, C. (2020). Digital platforms: Mapping the territory of new technologies to fight food waste. *British Food Journal*, 122(5), 1647–1669. doi:10.1108/BFJ-06-2019-0391
- Caplan, S. (1990). Using focus group methodology for ergonomic design. *Ergonomics*, 33(5), 527–533. doi:10.1080/00140139008927160
- Chao, G. K. (2021). *Applying Facial Emotion Recognition to Usability Evaluations to Reduce Analysis Time*. Cal Poly.
- Chen, L., & Pu, P. (2010). Eye-tracking study of user behavior in recommender interfaces. In *International conference on user modeling, adaptation, and personalization*. Springer. doi:10.1007/978-3-642-13470-8_35
- Cho, H., Powell, D., Pichon, A., Kuhns, L. M., Garofalo, R., & Schnall, R. (2019). Eye-tracking retrospective think-aloud as a novel approach for a usability evaluation. *International Journal of Medical Informatics*, 129, 366–373. doi:10.1016/j.ijmedinf.2019.07.010 PMID:31445278
- Chuka-Maduji, N. (2021). *On-Focus: A Web-Based Platform for Planning, Organizing, and Managing Focus Groups*.
- Clemmensen, T. (2009). Towards a theory of cultural usability: A comparison of ADA and CM-U theory. In *International Conference on Human Centered Design*. Springer. doi:10.1007/978-3-642-02806-9_48
- Constantine, L. L., & Lockwood, L. A. (1999). *Software for use: a practical guide to the models and methods of usage-centered design*. Pearson Education.
- Cook, A. S., & Meyer, B. (2017). Assessing leadership behavior with observational and sensor-based methods: A brief overview. *Handbook of methods in leadership research*.
- Cui, M., Wang, T., Pan, Z., & Ni, L. (2020). Definition of People with Impediments and Universality Evaluation of Public Service in Airport Travel Scenarios. In *The International Conference on Human-Computer Interaction*. Springer. doi:10.1007/978-3-030-49760-6_44
- Das, S., McEwan, T., & Douglas, D. (2008). Using eye-tracking to evaluate label alignment in online forms. In *Proceedings of the 5th Nordic conference on Human-computer interaction: building bridges*. ACM. doi:10.1145/1463160.1463217
- Davis, R., Gardner, J., & Schnall, R. (2020). A Review of Usability Evaluation Methods and Their Use for Testing eHealth HIV Interventions. *Current HIV/AIDS Reports*, 17(3), 1–16. doi:10.1007/s11904-020-00493-3 PMID:32390078

Dennerlein, J., Becker, T., Johnson, P., Reynolds, C., & Picard, R. W. (2003). Frustrating computer users increases exposure to physical factors. *Proceedings of the International Ergonomics Association*.

Dewi, P., Dantes, G., & Indrawan, G. (2020). User experience evaluation of e-report application using cognitive walkthrough (cw), heuristic evaluation (he) and user experience questionnaire (ueq). *Journal of Physics: Conference Series*. doi:10.1088/1742-6596/1516/1/012024

Díaz, J., Rusu, C., & Collazos, C. A. (2017). Experimental validation of a set of cultural-oriented usability heuristics: e-Commerce websites evaluation. *Computer Standards & Interfaces*, 50, 160–178. doi:10.1016/j.csi.2016.09.013

do Amaral, V., Ferreira, L. A., Aquino, P. T., & de Castro, M. C. F. (2013). EEG signal classification in usability experiments. In *2013 ISSNIP Biosignals and Biorobotics Conference: Biosignals and Robotics for Better and Safer Living (BRC)*. IEEE. doi:10.1109/BRC.2013.6487469

Ehmke, C., & Wilson, S. (2007). *Identifying web usability problems from eyetracking data*. Academic Press.

Elling, S., Lentz, L., & de Jong, M. (2011). Retrospective think-aloud method: Using eye movements as an extra cue for participants' verbalizations. *Proceedings of the SIGCHI conference on human factors in computing systems*. doi:10.1145/1978942.1979116

Falkowska, J., Sobiecki, J., & Pietrzak, M. (2016). *Eye tracking usability testing enhanced with EEG analysis*. Paper presented at the International Conference of Design, User Experience, and Usability. doi:10.1007/978-3-319-40409-7_38

Fan, M. (2019). *Leveraging Subtle Verbalization and Speech Patterns to Help Evaluators Identify Usability Problem Encounters in Concurrent Think-aloud Sessions*. University of Toronto.

Fitzpatrick, R. (1998). Strategies for evaluating software usability. *Methods (San Diego, Calif.)*, 353(1).

Foglia, P., Prete, C. A., & Zanda, M. (2008). *Relating GSR signals to traditional usability metrics: Case study with an anthropomorphic web assistant*. Paper presented at the 2008 IEEE Instrumentation and Measurement Technology Conference. doi:10.1109/IMTC.2008.4547339

Forçan, L. R., Abe, J. M., de Lima, L. A., & Nascimento, S. S. (2020). *Questionnaire Model for Paraconsistent Quality Assessment of Software Developed in Salesforce*. Paper presented at the IFIP International Conference on Advances in Production Management Systems.

Forster, Y., Hergeth, S., Naujoks, F., & Krems, J. F. (2018). How usability can save the day—methodological considerations for making automated driving a success story. *Proceedings of the 10th international conference on automotive user interfaces and interactive vehicular applications*. doi:10.1145/3239060.3239076

Friberg, A. (2017). Why continuous usability testing can and should be part of regular library activity—from a UX librarian's point of view. *Revy*, 40(1), 9–11.

Friedman, B., Khan, P. H. Jr, & Howe, D. C. (2000). Trust online. *Communications of the ACM*, 43(12), 34–40. doi:10.1145/355112.355120

Friedrichs, T., Borojeni, S. S., Heuten, W., Lüdtke, A., & Boll, S. (2016). PlatoonPal: User-Centered Development and Evaluation of an Assistance System for Heavy-Duty Truck Platooning. *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. doi:10.1145/3003715.3005449

Geisen, E., & Murphy, J. (2020). A compendium of web and mobile survey pretesting methods. *Advances in questionnaire design, development, evaluation and testing*, 287-314.

Goertzen, M., & Bakkalbasi, N. (2016). *Exploring academic e-book use: part II through focus groups and interviews*. Performance Measurement and Metrics.

Goh, K. N., Chen, Y. Y., Lai, F., Daud, S. C., Sivaji, A., & Soo, S.-T. (2013). *A comparison of usability testing methods for an e-commerce website: A case study on a Malaysia online gift shop*. Paper presented at the 2013 10th International Conference on Information Technology: New Generations. doi:10.1109/ITNG.2013.129

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

- Gong, R., & Elkerton, J. (1990). Designing minimal documentation using a GOMS model: A usability evaluation of an engineering approach. *Proceedings of the SIGCHI conference on Human factors in computing systems*. doi:10.1145/97243.97261
- Han, S. H., Yun, M. H., Kim, K.-J., & Kwahk, J. (2000). Evaluation of product usability: Development and validation of usability dimensions and design elements based on empirical models. *International Journal of Industrial Ergonomics*, 26(4), 477–488. doi:10.1016/S0169-8141(00)00019-6
- Han, S. H., Yun, M. H., Kwahk, J., & Hong, S. W. (2001). Usability of consumer electronic products. *International Journal of Industrial Ergonomics*, 28(3-4), 143–151. doi:10.1016/S0169-8141(01)00025-7
- Hartson, H. R., Andre, T. S., & Williges, R. C. (2001). Criteria for evaluating usability evaluation methods. *International Journal of Human-Computer Interaction*, 13(4), 373–410. doi:10.1207/S15327590IJHC1304_03
- Hasan, L., Morris, A., & Proberts, S. (2009). *Using Google Analytics to evaluate the usability of e-commerce sites*. Paper presented at the International Conference on Human Centered Design. doi:10.1007/978-3-642-02806-9_81
- Hasibuan, D., Santoso, H., Yunita, A., & Rahmah, A. (2020). An Indonesian Adaptation of the E-Learning Usability Scale. *Journal of Physics: Conference Series*. doi:10.1088/1742-6596/1566/1/012051
- Hassan, I. (2020). The Graphical Experience: User Interface Design Approach Based on User-Centered Design to Support Usability. In *Internet of Things—Applications and Future* (pp. 377–399). Springer. doi:10.1007/978-981-15-3075-3_26
- Hoffmann, E. R. (2013). Which version/variation of Fitts' law? A critique of information-theory models. *Journal of Motor Behavior*, 45(3), 205–215. doi:10.1080/00222895.2013.778815 PMID:23581725
- Holden, R. J. (2020). A simplified system usability scale (SUS) for cognitively impaired and older adults. *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*. doi:10.1177/2327857920091021
- Holdnack J. A. Brennan P. F. (2021). Usability and Effectiveness of Immersive Virtual Grocery Shopping for Assessing Cognitive Fatigue in Healthy Controls: Protocol for a Randomized Controlled Trial (Preprint). *JMIR Research Protocols*. 10.2196/preprints.28073
- Hornbæk, K. (2010). Dogmas in the assessment of usability evaluation methods. *Behaviour & Information Technology*, 29(1), 97–111. doi:10.1080/01449290801939400
- Horton, E. S. (2011). *Enhanced Browsing of Digital Archival Collections: A Pilot Study*. Academic Press.
- Hsu, L., & Chen, Y. (2019). Music and wine tasting: An experimental neuromarketing study. *British Food Journal*, 122(8), 2725–2737. doi:10.1108/BFJ-06-2019-0434
- Hwang, Y. M., & Lee, K. C. (2020). An eye-tracking paradigm to explore the effect of online consumers' emotion on their visual behaviour between desktop screen and mobile screen. *Behaviour & Information Technology*, 1–12.
- Ishaq, K., Zin, N. A. M., Rosdi, F., Abid, A., & Ali, Q. (2020). Usefulness of mobile assisted language learning in primary education. *International Journal of Advanced Computer Science and Applications*, 11(1), 383–395. doi:10.14569/IJACSA.2020.0110148
- Ivory, M. Y., & Hearst, M. A. (2001). The state of the art in automating usability evaluation of user interfaces. *ACM Computing Surveys*, 33(4), 470–516. doi:10.1145/503112.503114
- Ji, X., & Rau, P.-L. P. (2019). A comparison of three think-aloud protocols used to evaluate a voice intelligent agent that expresses emotions. *Behaviour & Information Technology*, 38(4), 375–383. doi:10.1080/0144929X.2018.1535621
- Jokela, T., Iivari, N., Tornberg, V., & Electro, P. (2004). Using the ISO 9241-11 definition of usability in requirements determination: case studies. In *Proceedings of HCI2004: Design for life, the 18th British HCI group annual conference*. Leeds Metropolitan University.
- Juarez, D., Tur-Viñes, V., & Mengual, A. (2020). Neuromarketing Applied to Educational Toy Packaging. *Frontiers in Psychology*, 11, 2077. doi:10.3389/fpsyg.2020.02077 PMID:32982857

- Karweit, N. (1988). Time-on-task: The second time around. *NASSP Bulletin*, 72(505), 31–39. doi:10.1177/019263658807250507
- Keinan, G. (1987). Decision making under stress: Scanning of alternatives under controllable and uncontrollable threats. *Journal of Personality and Social Psychology*, 52(3), 639–644. doi:10.1037/0022-3514.52.3.639 PMID:3572731
- Keinan, G., Friedland, N., & Ben-Porath, Y. (1987). Decision making under stress: Scanning of alternatives under physical threat. *Acta Psychologica*, 64(3), 219–228. doi:10.1016/0001-6918(87)90008-4
- Khosla, R., Damiani, E., & Grosky, W. (2003). Why Human-Centered e-Business? In *Human-centered e-business* (pp. 1-12). Springer. doi:10.1007/978-1-4615-0445-0_1
- Kieras, D. (1997). A guide to GOMS model usability evaluation using NGOMSL. In *Handbook of human-computer interaction* (pp. 733–766). Elsevier. doi:10.1016/B978-044481862-1.50097-2
- Kieras, D. E. (1999). A guide to GOMS model usability evaluation using GOMSL and GLEAN3. University of Michigan.
- Kirakowski, J. (1998). *Website Analysis and Measurement Inventory*. Academic Press.
- Kirakowski, J., & Cierlik, B. (1998). Measuring the usability of web sites. *Proceedings of the Human Factors and Ergonomics Society annual meeting*.
- Klaib, A. F., Alsrehin, N. O., Melhem, W. Y., Bashtawi, H. O., & Magableh, A. A. (2021). Eye tracking algorithms, techniques, tools, and applications with an emphasis on machine learning and Internet of Things technologies. *Expert Systems with Applications*, 166, 114037. doi:10.1016/j.eswa.2020.114037
- Kramer, J., Noronha, S., & Vergo, J. (2000). A user-centered design approach to personalization. *Communications of the ACM*, 43(8), 44–48. doi:10.1145/345124.345139
- Krebs, C., Falkner, M., Niklaus, J., Persello, L., Klöppel, S., Nef, T., & Urwyler, P. (2021). Application of Eye Tracking in Puzzle Games for Adjunct Cognitive Markers: Pilot Observational Study in Older Adults. *JMIR Serious Games*, 9(1), e24151. doi:10.2196/24151 PMID:33749607
- Kumar, R., Khan, A. I., Abushark, Y. B., Alam, M. M., Agrawal, A., & Khan, R. A. (2020). An integrated approach of fuzzy logic, AHP and TOPSIS for estimating usable-security of web applications. *IEEE Access : Practical Innovations, Open Solutions*, 8, 50944–50957. doi:10.1109/ACCESS.2020.2970245
- Kwahk, J., & Han, S. H. (2002). A methodology for evaluating the usability of audiovisual consumer electronic products. *Applied Ergonomics*, 33(5), 419–431. doi:10.1016/S0003-6870(02)00034-0 PMID:12236651
- Lagha, R. R., Burningham, Z., Sauer, B. C., Leng, J., Peters, C., Huynh, T., & Kramer, B. J. et al. (2020). Usability Testing a Potentially Inappropriate Medication Dashboard: A Core Component of the Dashboard Development Process. *Applied Clinical Informatics*, 11(04), 528–534. doi:10.1055/s-0040-1714693 PMID:32785904
- Lattie, E. G., Bass, M., Garcia, S. F., Phillips, S. M., Moreno, P. I., Flores, A. M., & Penedo, F. J. et al. (2020). Optimizing health information technologies for symptom management in cancer patients and survivors: Usability evaluation. *JMIR Formative Research*, 4(9), e18412. doi:10.2196/18412 PMID:32955450
- Lewis, C., & Wharton, C. (1997). Cognitive walkthroughs. In *Handbook of human-computer interaction* (pp. 717–732). Elsevier. doi:10.1016/B978-044481862-1.50096-0
- Lewis, J. R. (1991). Psychometric evaluation of an after-scenario questionnaire for computer usability studies: The ASQ. *ACM Sigchi Bulletin*, 23(1), 78–81. doi:10.1145/122672.122692
- Lewis, K. M., & Hepburn, P. (2010). Open card sorting and factor analysis: A usability case study. *The Electronic Library*, 28(3), 401–416. doi:10.1108/02640471011051981
- Liapis, A., Katsanos, C., Karousos, N., Sotiropoulos, D., Xenos, M., & Orphanoudakis, T. (2020). *Stress Heatmaps: A Fuzzy-Based Approach that Uses Physiological Signals*. Paper presented at the International Conference on Human-Computer Interaction. doi:10.1007/978-3-030-49757-6_19
- Likierman, M. (2006). Unconscious experience: Relational perspectives. *Psychoanalytic Dialogues*, 16(4), 365–376.

- Lin, T., Omata, M., Hu, W., & Imamiya, A. (2005). Do physiological data relate to traditional usability indexes? *Proceedings of the 17th Australia conference on computer-human interaction: Citizens online: Considerations for today and the future*.
- Liu, Y., Osvalder, A.-L., & Dahlman, S. (2005). Exploring user background settings in cognitive walkthrough evaluation of medical prototype interfaces: A case study. *International Journal of Industrial Ergonomics*, 35(4), 379–390. doi:10.1016/j.ergon.2004.10.004
- Luan, J., Yao, Z., Zhao, F., & Liu, H. (2016). Search product and experience product online reviews: An eye-tracking study on consumers' review search behavior. *Computers in Human Behavior*, 65, 420–430. doi:10.1016/j.chb.2016.08.037
- Lund, A. M. (1997). Expert ratings of usability maxims. *Ergonomics in Design*, 5(3), 15–20. doi:10.1177/106480469700500305
- Madan, A., & Dubey, S. K. (2012). Usability evaluation methods: A literature review. *International Journal of Engineering Science and Technology*, 4(2), 590–599.
- Malehmir, F., Maeen, M., & Jahangir, M. R. (2017). A study on the Interaction of Motivations and Online Shopping Experience in E-Commerce Success in Digikala Company. *International Journal of Scientific Study*, 5(5).
- Märtin, C., Bissinger, B. C., & Asta, P. (2021). Optimizing the digital customer journey—Improving user experience by exploiting emotions, personas and situations for individualized user interface adaptations. *Journal of Consumer Behaviour*, cb.1964. doi:10.1002/cb.1964
- McCloskey, M. L., Johnson, S. L., Bekelman, T. A., Martin, C. K., & Bellows, L. L. (2019). Beyond nutrient intake: use of digital food photography methodology to examine family dinnertime. *Journal of Nutrition Education and Behavior*, 51(5), 547-555.
- McNamara, N., & Kirakowski, J. (2006). Functionality, usability, and user experience: three areas of concern. *Interactions*, 13(6), 26-28.
- Modi, N., & Singh, J. (2022). Understanding Online Consumer Behavior at E-commerce Portals Using Eye-Gaze Tracking. *International Journal of Human-Computer Interaction*, 39(4), 721–742. doi:10.1080/10447318.2022.2047318
- Mol, M., van Schaik, A., Dozeman, E., Ruwaard, J., Vis, C., Ebert, D. D., & Mora, T. et al. (2020). Dimensionality of the system usability scale among professionals using internet-based interventions for depression: A confirmatory factor analysis. *BMC Psychiatry*, 20(1), 1–10. doi:10.1186/s12888-020-02627-8 PMID:32398111
- Monica, . B., Iuliana, C., & Mihai, . (2019). Studying the User Experience in Online Banking Services: An Eye-Tracking Application. *Studies in Business and Economics*, 14(2), 193–208. doi:10.2478/sbe-2019-0034
- Nielsen, J. (1994a). *Usability engineering*. Morgan Kaufmann.
- Nielsen, J. (1994b). *Usability inspection methods*. Paper presented at the Conference companion on Human factors in computing systems. doi:10.1145/259963.260531
- Nielsen, J. (2011). *Top 10 mistakes in web design*. Nielsen Norman Group.
- Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. *Proceedings of the SIGCHI conference on Human factors in computing systems*.
- Nielsen, J., & Pernice, K. (2010). *Eyetracking web usability*. New Riders.
- Nilashi, M., Samad, S., Ahmadi, N., Ahani, A., Abumalloh, R. A., Asadi, S., & Yadegaridehkordi, E. et al. (2020). Neuromarketing: A Review of Research and Implications for Marketing. *Journal of Soft Computing and Decision Support Systems*, 7(2), 23–31.
- Niu, Y., Zhou, T., & Bai, L. (2021). Research on color coding of fighter jet head-up display key information elements in air–sea flight environment based on eye-tracking technology. *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*.
- Olsen, A. (2012). *The Tobii IVT Fixation Filter Algorithm description*. Academic Press.

- Page, T. (2011). Interaction and usability considerations in the design of mobile phones. *Journal of Desert Research*, 9(3), 281–300. doi:10.1504/JDR.2011.041394
- Paknejad, F., Mosaddad, S. H., & Sadeghi Naeini, H. (2021). Purchasing and Consumption Modification Among Iranians Throughout Gamification. *International Journal of Industrial Engineering & Production Research*, 32(1), 121–132.
- Papavlasopoulou, S., Sharma, K., & Giannakos, M. N. (2020). Coding activities for children: Coupling eye-tracking with qualitative data to investigate gender differences. *Computers in Human Behavior*, 105, 105939. doi:10.1016/j.chb.2019.03.003
- Petrie, H., & Kheir, O. (2007). The relationship between accessibility and usability of websites. *Proceedings of the SIGCHI conference on Human factors in computing systems*. doi:10.1145/1240624.1240688
- Pittig, A., Hoyer, J., & Noack, R. (2021). Smart-Glass Guided Exposure for Anxiety Disorders: A Proof-of-Concept Study. *Cognitive and Behavioral Practice*, 28(3), 364–378. doi:10.1016/j.cbpra.2020.12.003
- Porcelli, A. J., & Delgado, M. R. (2009). Acute stress modulates risk taking in financial decision making. *Psychological Science*, 20(3), 278–283. doi:10.1111/j.1467-9280.2009.02288.x PMID:19207694
- Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., & Carey, T. (1994). *Human-Computer Interaction*. Addison-Wesley.
- Rauf, I., Troubitsyna, E., & Porres, I. (2019). A systematic mapping study of API usability evaluation methods. *Computer Science Review*, 33, 49–68. doi:10.1016/j.cosrev.2019.05.001
- Rawnaque, F. S., Rahman, K. M., Anwar, S. F., Vaidyanathan, R., Chau, T., Sarker, F., & Al Mamun, K. A. (2020). Technological advancements and opportunities in Neuromarketing: A systematic review. *Brain Informatics*, 7(1), 1–19. doi:10.1186/s40708-020-00109-x PMID:32955675
- Reynolds, C. J. (2005). *Adversarial uses of affective computing and ethical implications*. Massachusetts Institute of Technology.
- Rice, A. D., & Lartigue, J. W. (2014). Touch-level model (TLM) evolving KLM-GOMS for touchscreen and mobile devices. *Proceedings of the 2014 ACM Southeast regional conference*. doi:10.1145/2638404.2638532
- Rich, A., & McGee, M. (2004). Expected usability magnitude estimation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*.
- Rieman, J., Franzke, M., & Redmiles, D. (1995). *Usability evaluation with the cognitive walkthrough*. Paper presented at the Conference companion on Human factors in computing systems.
- Rodrigues, I. P., de Borba Campos, M., & Zorzo, A. F. (2017). *Usability evaluation of domain-specific languages: A systematic literature review*. Paper presented at the International Conference on Human-Computer Interaction.
- Roman, L. C., Ancker, J. S., Johnson, S. B., & Senathirajah, Y. (2017). Navigation in the electronic health record: A review of the safety and usability literature. *Journal of Biomedical Informatics*, 67, 69–79. doi:10.1016/j.jbi.2017.01.005 PMID:28088527
- Sadeghi Naeini, H., Dalal, K., Mosaddad, H., & Karupiah, K. (2018). Economic effectiveness of Ergonomics interventions. *International Journal of Industrial Engineering & Production Research*, 29(3), 261–276. doi:10.22068/ijiepr.29.3.261
- Sargezeh, B. A., Tavakoli, N., & Daliri, M. R. (2019). Gender-based eye movement differences in passive indoor picture viewing: An eye-tracking study. *Physiology & Behavior*, 206, 43–50. doi:10.1016/j.physbeh.2019.03.023 PMID:30922820
- Sauro, J. (2015). SUPR-Q: A comprehensive measure of the quality of the website user experience. *Journal of Usability Studies*, 10(2).
- Sauro, J., & Dumas, J. S. (2009). Comparison of three one-question, post-task usability questionnaires. *Proceedings of the SIGCHI conference on human factors in computing systems*. doi:10.1145/1518701.1518946
- Sauro, J., & Lewis, J. R. (2016). *Quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann.

- Savela-Huovinen, U., Toom, A., Knaapila, A., & Muukkonen, H. (2021). Sensory professionals' perspective on the possibilities of using facial expression analysis in sensory and consumer research. *Food Science & Nutrition*, 9(8), 4254–4265. doi:10.1002/fsn3.2393 PMID:34401076
- Schmidt, T., Wittmann, V., & Wolff, C. (2019). The Influence of Participants' Personality on Quantitative and Qualitative Metrics in Usability Testing. *Proceedings of Mensch und Computer, 2019*, 115–126. doi:10.1145/3340764.3340787
- Sneiderman, B. (1992). *Designing the user interface: Strategies for effective human-computer interaction*. Addison-Wesley.
- Seffah, A., Donyaee, M., Kline, R. B., & Padda, H. K. (2006). Usability measurement and metrics: A consolidated model. *Software Quality Journal*, 14(2), 159–178. doi:10.1007/s11219-006-7600-8
- Shackel, B. (2009). Usability–Context, framework, definition, design and evaluation. *Interacting with Computers*, 21(5-6), 339–346. doi:10.1016/j.intcom.2009.04.007
- Shneiderman, B. (1986). Eight golden rules of interface design. *Disponibile en*, 172.
- Singh, T., Malik, S., & Sarkar, D. (2016). *E-commerce website quality assessment based on usability*. Paper presented at the 2016 International Conference on Computing, Communication and Automation (ICCCA). doi:10.1109/CCAA.2016.7813698
- Siroker, D., & Koomen, P. (2013). *A/B testing: The most powerful way to turn clicks into customers*. John Wiley & Sons.
- Sivaji, A., Downe, A. G., Mazlan, M. F., Soo, S.-T., & Abdullah, A. (2011). *Importance of incorporating fundamental usability with social & trust elements for e-commerce website*. Paper presented at the 2011 International Conference on Business, Engineering and Industrial Applications. doi:10.1109/ICBEIA.2011.5994248
- Snyder, M. E., Adeoye-Olatunde, O. A., Gernant, S. A., DiIulio, J., Jaynes, H. A., Doucette, W. R., & Russ-Jara, A. L. (2021). A user-centered evaluation of medication therapy management alerts for community pharmacists: Recommendations to improve usability and usefulness. *Research in Social & Administrative Pharmacy*, 17(8), 1433–1443. doi:10.1016/j.sapharm.2020.10.015 PMID:33250363
- Soranzo, A., & Cooksey, D. (2015). Testing taxonomies: Beyond card sorting. *Bulletin of the Association for Information Science and Technology*, 41(5), 34–39. doi:10.1002/bult.2015.1720410509
- Sripathi, V., & Sandru, V. (2013). Effective usability testing–knowledge of user centered design is a key requirement. *International Journal of Emerging Technology and Advanced Engineering*, 3(1), 627–635.
- Thoma, V., & Dodd, J. (2019). Web Usability and Eyetracking. In *Eye Movement Research* (pp. 883–927). Springer. doi:10.1007/978-3-030-20085-5_21
- Tognazzini, B. (2014). *First principles of interaction design (revised & expanded)*. Ask TOG: Interaction design solutions for the real world. <https://asktog.com/atc/principles-of-interaction-design>
- Tomeo, M. L. (2012). Adding users to the Website design process. *Public Services Quarterly*, 8(4), 350–358. doi:10.1080/15228959.2012.730414
- Tomitsch, M., Janssen, A., Curwood, J. S., & Thomson, K. (2020). Insights into the interrelationship between the design process and user practice: A case study on user engagement in a platform for professional learning. *Enterprise Information Systems*, 1–16.
- Trimmel, M., Meixner-Pendleton, M., & Haring, S. (2003). Stress response caused by system response time when searching for information on the Internet. *Human Factors*, 45(4), 615–622. doi:10.1518/hfes.45.4.615.27084 PMID:15055458
- Tullis, T. S., & Stetson, J. N. (2004). *A comparison of questionnaires for assessing website usability*. Paper presented at the Usability professional association conference.
- Tupikovskaja-Omovie, Z., & Tyler, D. J. (2020). *Mobile Consumer Behavior in Fashion m-Retail: An Eye Tracking Study to Understand Gender Differences*. Paper presented at the ACM Symposium on Eye Tracking Research and Applications. doi:10.1145/3379157.3391305

Tupikovskaja-Omovie, Z., & Tyler, D. J. (2021). Experienced versus inexperienced mobile users: eye tracking fashion consumers' shopping behaviour on smartphones. *International Journal of Fashion Design, Technology and Education*, 1-9.

Unrau, R., & Kray, C. (2019). Usability evaluation for geographic information systems: A systematic literature review. *International Journal of Geographical Information Science*, 33(4), 645–665. doi:10.1080/13658816.2018.1554813

Veilleux, M., Sénécal, S., Demolin, B., Bouvier, F., Di Fabio, M.-L., Coursaris, C., & Léger, P.-M. (2020). *Visualizing a User's Cognitive and Emotional Journeys: A Fintech Case*. Paper presented at the International Conference on Human-Computer Interaction. doi:10.1007/978-3-030-49713-2_38

Veto, P., Thomas, P., Alexander, P., Wemyss, T., & Mollon, J. (2020). 'The last channel': vision at the temporal margin of the field. *Proceedings of the Royal Society B*, 287(1927), 20200607. doi:10.1098/rspb.2020.0607

Villa, R., & Monzón, A. (2021). Mobility Restrictions and E-Commerce: Holistic Balance in Madrid Centre during COVID-19 Lockdown. *Economies*, 9(2), 57. doi:10.3390/economies9020057

Vlachogianni, P., & Tselios, N. (2021). Perceived usability evaluation of educational technology using the System Usability Scale (SUS): A systematic review. *Journal of Research on Technology in Education*, 1–18.

Wang, C., Terken, J., & Hu, J. (2014). "Liking" other Drivers' Behavior while Driving. *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. doi:10.1145/2667239.2667424

Ward, R. D., & Marsden, P. H. (2003). Physiological responses to different web page designs. *International Journal of Human-Computer Studies*, 59(1-2), 199–212. doi:10.1016/S1071-5819(03)00019-3

Wenzel, M. A., Schultze-Kraft, R., Meinecke, F. C., Cardinaux, F., Kemp, T., Müller, K.-R., & Blankertz, B. et al. (2015). EEG-based usability assessment of 3D shutter glasses. *Journal of Neural Engineering*, 13(1), 016003. doi:10.1088/1741-2560/13/1/016003 PMID:26644071

Wilson, C. (2013). *Credible checklists and quality questionnaires: A user-centered design method*. Newnes.

Zardari, B. A., Hussain, Z., Arain, A. A., Rizvi, W. H., & Vighio, M. S. (2021). QUEST e-learning portal: Applying heuristic evaluation, usability testing and eye tracking. *Universal Access in the Information Society*, 20(3), 531–543. doi:10.1007/s10209-020-00774-z

Zhang, W., & Liu, H. (2017). Toward a reliable collection of eye-tracking data for image quality research: Challenges, solutions, and applications. *IEEE Transactions on Image Processing*, 26(5), 2424–2437. doi:10.1109/TIP.2017.2681424 PMID:28362586

Zimmerman, D. E., & Akerelrea, C. (2002). A group card sorting methodology for developing informational web sites. *Proceedings. IEEE International Professional Communication Conference*. doi:10.1109/IPCC.2002.1049127

Ali Bonyadi Naeini is a faculty member of the Business Department of Iran University of Science and Technology (IUST). He has taught several courses in the fields of marketing management, marketing research, and research design & method so far. He has also presented several conference papers and scholarly journal papers mostly focused on quantitative decision-making techniques.

Ali Golbazi Mahdipour received his B.Sc. degree in electrical engineering from the University of Tabriz, Tabriz, in 2015 and his M.Sc. degree in business management from the Iran University of Science and Technology (IUST) in 2020. He is currently a full-time research fellow in the Neurobusiness Lab. at IUST. His main research interests include neuroeconomics, consumer neuroscience, and computational and cognitive neuroscience.

Rasam Dorri received his B.Sc. degree in Civil Engineering from the Azad University, Tehran, and his M.Sc. degree in business administration from the Iran University of Science and Technology (IUST) in 2022. He is currently a part-time research fellow in the Neurobusiness Lab. at IUST. His main research interests include consumer behavior, consumer neuroscience, and computational and cognitive neuroscience.