How quickly can we predict users' ratings on aesthetic evaluations of websites? Employing machine learning on eye-tracking data

Ilias O. Pappas[1, 2 \[0000-0001-7528-3488\]], Kshitij Sharma[2 \[0000-0003-3364-637X\]], Patrick Mikalef [2 \[0000-0002-6788-2277\]] and Michail N. Giannakos[2 \[0000-0002-8016-6208\]]

[1] University of Agder, Kristiansand, NO-4639, Norway
[2] Norwegian University of Science and Technology, Trondheim, NO-7491, Norway
ilias.pappas@uia.no; ilpappas@ntnu.no

Abstract. This study examines how quickly we can predict users' ratings on visual aesthetics in terms of simplicity, diversity, colorfulness, craftsmanship. To predict users' ratings, first we capture gaze behavior while looking at high, neutral, and low visually appealing websites, followed by a survey regarding user perceptions on visual aesthetics towards the same websites. We conduct an experiment with 23 experienced users in online shopping, capture gaze behavior and through employing machine learning we examine how fast we can accurately predict their ratings. The findings show that after 25 seconds we can predict ratings with an error rate ranging from 9% to 11% depending on which facet of visual aesthetic is examined. Furthermore, within the first 15 seconds we can have a good and sufficient prediction for simplicity and colorfulness, with error rates 11% and 12% respectively. For diversity and craftsmanship, 20 seconds are needed to get a good and sufficient prediction similar to the one from 25 seconds. The findings indicate that we need more than 10 seconds of viewing time to be able to accurately capture perceptions on visual aesthetics. The study contributes by offering new ways for designing systems that will take into account users' gaze behavior in an unobtrusive manner and will be able inform researchers and designers about their perceptions of visual aesthetics.

Keywords: Eye-tracking, Machine Learning, E-commerce, Aesthetics, Design, Artificial Intelligence.

1 Introduction

Website and application development have largely focused on good visual design that can support users' interaction with the website and improve their online experience. To this end, previous work has studied several website design characteristics, including their effectiveness and visual appeal [1-4]. Daily, we visit numerous websites or use their applications, and our evaluations are affected by their usability, how they look and feel, as well as their content and service quality [5]. Since website design will influence user experience, designers and the companies they work for strive to offer aesthetically

pleasing websites of high quality. The majority of the studies in the area employ subjective methods (e.g., data collection using questionnaires), while some utilize more objective approaches by collecting physiological data such as fMRI and eye-tracking [6, 7].

The advancement of technology and machine learning techniques allows the implementation of Artificial Intelligence (AI) applications in many fields, including Human Computer Interaction [8] or image aesthetics [9]. Recently, early attempts have been proposed to employ machine learning for assessing aesthetics of websites [10]. Furthermore, previous studies have assessed website aesthetics using eye-tracking techniques to detect areas of interest while looking at a website, mainly by using heatmaps [11-13]. Going beyond heatmaps, eye-tracking techniques can offer deeper insight into users' gaze behavior as they can also capture how one's eyes move during the whole period, including for example speed of eye movement, a change in direction, switching between two spots [14, 15].

Users' first impression is very important when they visit new websites, or they visit re-designed websites. Earlier studies suggest that first impressions are formed very fast when users visit at a website, in the very first seconds [16]. Thus, when taking design decisions different aspects that grab one's attention are more likely to influence their perceptions on visual aesthetics. However, as users become more experienced their perceptions regarding visual aesthetics are likely to evolve as well [15]. Indeed, eye-tracking data show that the more a user fixates on certain elements in a website the more likely it is to find it less appealing [14], which may occur if users start looking into more details as they look at the website. This raises the question if and how fast we can predict users' perceptions on visual aesthetics while they are looking at a website. Thus, we formulate the following research question: *RQ: How fast can we predict users' ratings when judging visual aesthetics in websites in terms of simplicity, diversity, colorfulness, and craftsmanship?*

To address our research question, first we measure how users' look at a website using eye-tracking along with their perceptions on visual aesthetics via a questionnaire. Then, we employ machine learning techniques to predict users' ratings from the gaze behavior and test how fast we can get an accurate prediction with a low error. We conducted an experiment in which users were asked to look at websites and then rate them in terms of simplicity, diversity, colorfulness, craftsmanship. The experiment included websites that ranged from low, neutral, and high visually appealing. We contribute by showing that when examining users' first impressions towards a website more than a few seconds are needed to get an accurate and valuable estimate on their perceptions towards the website's visual appeal. This is a first step towards designing and developing AI applications that will track users' eyes only for a specific amount of time (i.e., the time needed to get an accurate estimate) and inform designers about users' perceptions on visual aesthetics.

The paper is organized as follows. Section 2 presents the background of the study, along with related work on website visual aesthetics. Section 3 presents the details on the experiment, and the methodology that was employed. Section 4 presents the findings, and section 5 provides a discussion along with implications and suggestions for future work.

2   Background and conceptual model

2.1   Visual aesthetics of websites

Previous studies examining user experience highlight several factors that may influence users' evaluations towards visual aesthetics of websites [1-3, 17]. Capturing such evaluations for websites is highly important for website designers and can be done by examining screen design factors and layout elements which are related with users' perceptions of visual aesthetics [3]. To measure perceptions on visual aesthetics studies commonly use questionnaires which introduce a subjective take on the matter. Visual aesthetics can be divided into two basic types [4]. First, there are classical aesthetics that describe websites that are clear, well organized, and pleasant. Then, there are expressive aesthetics which refer to websites that are creative, colorful, and original.

Additionally, a more precise description of visual aesthetics has been proposed which offers a measure that represents visual aesthetics in a more holistic manner [2]. In detail, this measure includes four dimensions, or facets of visual aesthetics, which are simplicity, diversity, colorfulness, and craftsmanship. Simplicity, part of the classical aesthetics, refers to website design regarding its unity, clarity, orderliness and balance. Next, diversity, part of the expressive aesthetics, refers to a website design having a dynamic, novel, and creative design. For a more holistic representation of visual aesthetics the measure was extended by adding colorfulness and craftsmanship, adopted from Lavie and Tractinsky [4]. The unique effect of colors as well as the skillful and coherent integration of all design dimensions can be evaluated [1, 2]. The present study adopts the four facets of visual aesthetics as they are considered to be more detailed by examining more aspects of aesthetics [2]. This approach enables a better and deeper understanding of perceived visual aesthetics.

2.2   Visual aesthetics and eye-tracking

HCI studies that examine visual aesthetics of e-commerce websites measure users' gaze with different eye tracking techniques [11, 13], offering a better understanding of user behavior when visiting e-commerce websites. Indeed, we can collect data on where exactly the user looks when browsing a website and also how they look at it. The manner we look at an object is related with how we process and think about that object [18]. Our aesthetical preferences can be predicted with the derivatives of the stimulus properties [11, 19, 20]. Furthermore, based on the eye-mind hypothesis [18], what a user looks at indicates their cognitive process. These two notions can be simply described as *"what you perceive is what you see"* and *"what you see is what you process"*. Thus, cognitive load theory has been used to explain gaze behavior and how we process what we see [21], supporting the eye-mind hypothesis [18]. Consequently, what users perceive and what they process is also linked with their gaze.

By being able to capture measurable qualities of an object or a stimulus we can explain users' perceptions on visual aesthetics [22], thus here we employ eye-tracking techniques to examine how specific functions can be used to predict users' ratings on visual appeal. In detail, we use eye-tracking measures and compute their respective

4

features as a proxy for users perceived visual aesthetics, to examine how looking at websites explains perceptions of visual appeal. Previous studies have employed eye-tracking techniques to predict website design characteristics and users' visual attention [11, 13, 21, 23]. Nonetheless, previous studies mainly focus on using heat maps to explain eye behavior, which although is useful to show where the user focused, but it does not provide any information on how his or her eyes moved while looking at the website and interacting with it. Here we combine information from gaze behavior and questionnaires to predict users' ratings and extend previous studies by examining how fast can we get an accurate prediction.

3   Research Methodology

3.1   Context

Building on previous work [11, 12] the homepages of some popular e-commerce sites are used to examine their visual appeal. For this, we use the top-100 popular websites, based on the ranks on Alexa.com (i.e., provider of global traffic estimates in different categories). Fifty e-commerce websites were selected randomly, and 10 HCI experts were asked to evaluate them in terms of visual appeal on a 7-point scale. Amazon websites were removed from the list as Alexa is owned by Amazon to avoid any potential bias. Based on the median score from experts' ratings, nine websites were chosen, creating 3 categories, that is high, neutral and low visually appealing, with 3 websites each. Figure 1 shows examples from each category. A key distinction among the homepages of these websites is the photo-text ration, a main design characteristic in visually appealing websites [2, 11]. Since the first impression of a webpage is formed right away, as soon as one visits it, [16], investigating users' eye-tracking behavior on those very first seconds that they look at a webpage is important and can prove beneficial in explaining how behavior relates to main factors of visual aesthetics and design.
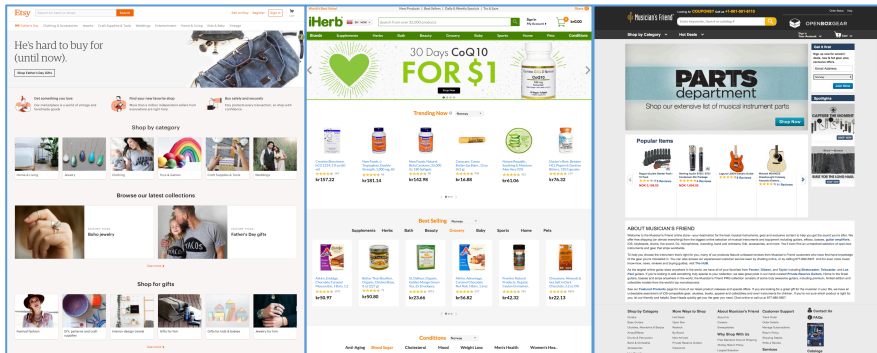

Fig. 1. High, Neutral, Low Appealing e-commerce websites (from left to right)

3.2   Sample and Procedure

Twenty-three users participated in this experiment (10 females, 13 males, average age 27.5 years, SD 7.15 years), with prior experience with e-commerce. Email and social media were used to find participants. We conducted the experiment at a lab space of a large university in Norway with 5 participants at a time.

We gathered information regarding age, gender, and internet experience, while we explained the objective of the study. To calibrate the eye tracker the participant watches a dot that moves to the four corners and center of the screen. The participants were instructed to look at each website for at least ten seconds. Then they were given a questionnaire which used a seven-point scale to capture their appeal ratings regarding the four facets of visual aesthetics. They viewed on average each webpage for about 25 seconds. Also, they were told not to base their ratings of visual appeal on the content of the page. We did not give them specific tasks because it can influence the way they look at the website and their focus [21], but they were instructed to look at the homepage freely. Participants had not purchased from the e-commerce websites before. Also, they controlled the mouse, allowing them to browse freely and decide on their own when to proceed to the next website.

3.3   Eye-tracking device

We used 4 SMI RED 250 and 1 TOBII mobile eye-tracker which look like glasses. They have a sampling rate of 60Hz, that is considered appropriate in usability studies [24]. The frames of the eye-tracker contain two cameras that record each eye of the participant; and an array of infrared LEDs which are reflected by the cornea. There is a third camera that records the field of view of the participants. Then the software calculates the position of the gaze on the field of view video. Finally, we correct for head movements by employing fiducial markers placed on the computer screen.

3.4   Measurements and feature extraction

For the visual aesthetics, we study simplicity, diversity, colorfulness, and craftsmanship [2]. Table 1 shows the actual and theoretical limits (min, max values) for the 4 facets of visual aesthetics. Minimum and maximum values are computed based on the 7-point scale and the fact that some items are reverse coded. Their items are presented in the Appendix.

Visual aesthetics	Min - Max	Theoretical Min - Max
Simplicity	2, 13	-11, 19
Diversity	-3, 11	-11, 19
Colorfulness	-4, 8	-12, 12
Craftmanship	-5, 8	-12, 12

Table 1. Thresholds for visual aesthetics

Eye tracking data offer the mean, variance, minimum, maximum and median of several parameters, such as pupil diameters, fixation details, saccade details, blink details,

and event statistics. Table 2 presents an overview of the extracted features along with the respective reference from the literature. In total, we extracted 31 features.

Eye-movement parameters	Features extracted	Source
Diameter	Pupil Diameter (mean, median, max, SD)	[25]
Fixation	Fixation duration (mean, median, max, SD)	[26]
	Fixation dispersion (mean, median, max, SD)	[27]
	Skewness of fixation duration histogram	[28]
Saccade	Ratio of forward saccades to total saccades	[29]
	Ratio of global and local saccades with a threshold on sac. vel.	[30]
	Skewness of saccade velocity histogram	[31]
	Saccade velocity (mean, median, max, SD)	[32]
	Saccade amplitude (mean, median, max, SD)	[33]
	Saccade duration (mean, median, max, SD)	[34]
Events	Number of fixations, number of saccades, fixation to saccade ratio	

Table 2. Eye-tracking features

3.5   Pipeline formation

In this paper, we present 4 pipelines. Each pipeline has 5 steps. First, we collect the data from 4 different interaction lengths: 1) 25 seconds, 2) 20 seconds, 3) 15 seconds and 10 seconds. Next, we compute the features given in the Table 1. Once the features are computed, we use a Random Forest regression algorithm to predict the four dependent variables. Next, we compare the prediction for all 4 different interaction lengths to examine how quickly we can get an accurate estimate.

Data → Pre-processing → Feature Extraction → Prediction using Random Forest Regression

Fig. 2. Pipeline formation

For each pipeline, we kept the data from 10% of the participants for the out of sample testing. We trained the random forest with the data from the remaining 90% participants using a 10-fold cross validation.
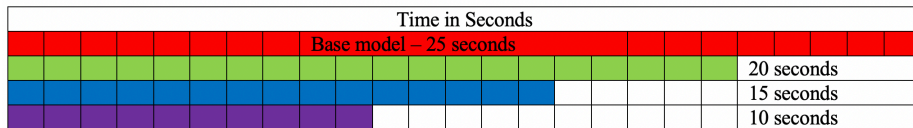

Fig. 3. Data collection for 4 interaction lengths

4   Findings

Table 3 depicts the prediction accuracy, that is the NRMSE (Normalized Root Mean Squared Error) prediction Random Forest regression algorithm by using the selected features (Table 2) for simplicity, diversity, colorfulness, and craftmanship. The

NRMSE value should be as low as possible. Tables 3 and 4 show which model can give a good prediction compared to the base model (25 seconds) and for which model the error rate is significantly different (Table 4) from the base model.

Table 3. Prediction accuracy for the 4 interaction lengths

Duration	Simplicity NRMSE (S.D.)	Diversity NRMSE (S.D.)	Colorfulness NRMSE (S.D.)	Craftsmanship NRMSE (S.D.)
25 seconds (base model)	0.09 (0.02)	0.09 (0.03)	0.10 (0.02)	0.11 (0.03)
20 seconds	0.10 (0.03)	**0.10 (0.04)**	0.11 (0.05)	**0.13 (0.05)**
15 seconds	**0.11 (0.05)**	0.12 (0.04)	**0.12 (0.05)**	0.14 (0.04)
10 seconds	0.16 (0.02)	0.15 (0.03)	0.17 (0.03)	0.16 (0.04)

Note: Normalized Root Mean Squared Error (NRMSE) prediction from Random Forest with the selected features (as presented in Table 2). In bold we show the Error is can be acceptable, compared to the base model, because it is **the one before the error being significantly different (based on T-test presented on Table 4)**

Table 4. Significant differences between the models

Duration	Simplicity	Diversity	Colorfulness	Craftsmanship
25 seconds (base model)	-	-	-	-
20 seconds	1.24 (0.22)	0.89 (0.37)	0.83 (0.43)	1.53 (0.13)
15 seconds	1.66 (0.10)	2.67 (0.01)**	1.66 (0.10)	2.68 (.01)**
10 seconds	11.06(.001)***	6.32 (0.001)***	8.68 (.001)***	4.47 (.001)***

Note: The difference of the base model with the other shorter in duration models. *** $p < 0.001$, ** $p < 0.01$

We can observe that in the case of simplicity, the model with 25 seconds of data has a 9% error while the models with the 20, 15 and 10 seconds of data have 10%, 11% and 16% errors, respectively. The model with 10 seconds of data is significantly worse than the model with 25 seconds of data. Hence, the chosen model is the one with the smallest error that is closest to the base model (model with 25 seconds of data) and uses the least amount of data. For simplicity the chosen model is the one with 15 seconds of data.

Next, in the case of diversity, the model with 25 seconds of data has a 9% error while the models with the 20, 15 and 10 seconds of data have 10%, 12% and 15% errors. The models with 15 and 10 seconds of data is significantly worse than the model with 25 seconds of data. Thus, the chosen model is the one with the smallest error that is closest to the base model (model with 25 seconds of data) and uses the least amount of data. For diversity the chosen model is the one with 20 seconds of data.

In the case of colorfulness, the model with 25 seconds of data has a 9% error while the models with the 20, 15 and 10 seconds of data have 11%, 12% and 17% errors. The model with 10 seconds of data is significantly worse than the model with 25 seconds of data. Similarly, the chosen model is the one with the smallest error that is closest to the base model (model with 25 seconds of data) and uses the least amount of data. For colorfulness the chosen model is the one with 15 seconds of data.

Finally, regarding craftsmanship, the model with 25 seconds of data has a 9% error while the models with the 20, 15 and 10 seconds of data have 13%, 14% and 16% errors. The models with 15 and 10 seconds of data are significantly worse than the model with 25 seconds of data. Similarly, the chosen model is the one with the smallest

error that is closest to the base model (model with 25 seconds of data) and uses the least amount of data. For craftsmanship the chosen model is the one with 20 seconds of data.

To offer more insight and details on what the NRMSE values show we should point out that they refer to the error on the range that our values get based on their theoretical limits (Table 1). For example, for simplicity the minimum and maximum theoretical limits are -11 and 19, respectively. Thus, the range of simplicity is 30. An 11% error on the range of 30 is basically a 3.3 deviance. Since, simplicity is calculated based on the sum of 5 items (Appendix), it means that per item, the 11% error translates to a 0.66 (i.e., 3.3/5) deviance on the Likert Scale. Thus, the accuracy of our predictions is quite high and can be deviant by approximately half point or less on the Likert scale.

5   Discussion

This study provides evidence on how quickly we can get accurate estimates of users' ratings regarding their perceptions of visual aesthetics measuring and analyzing their gaze behavior. Users perceptions are directly related to what they see, that is a visual stimulus, and the way that the stimulus changes will also influence users' perceptions. This has been described by two basic notions that are "what you perceive is what you see" [19, 20] and "what you see is what you process" [18]. This paper goes one step further from previous studies that predict users' ratings and answers the question on how quickly we can predict these ratings.

The findings show that for simplicity and colorfulness 15 seconds of gaze data offer a good prediction, similar to the one from 25 seconds of data. This means that when a user is looking at website, after 15 seconds they will have formed their first impression of it regarding simplicity and colorfulness and it will not change if they look at the website for up to half a minute. Furthermore, the findings show that 10 seconds are not sufficient to make a good enough prediction, compared to 25 seconds, and the error rate goes up to 16% and 17%, for simplicity and colorfulness. Previous studies suggest that our first impression is formed almost immediately [16] and users are instructed to look at websites for about 10 seconds in similar studies [11]. Similarly, in our study the participants were instructed to look at the websites at least for 10 seconds, but the findings show that by adding 5 more seconds (i.e., 15 in total) the prediction rate is similar to the prediction from 25 seconds.

In the case of diversity and craftmanship, the findings show that more data are needed for a good enough prediction. Thus, 20 seconds give a good prediction as 25 seconds of viewing. Similarly, to simplicity and colorfulness, 10 seconds are not enough to get a sufficient prediction as the error rates increase 15% and 16%, being significantly different from those with 25 seconds of data. We empirically show that in website aesthetic evaluation a user has to look at a website for 15 seconds or more depending on which facet of visual aesthetics is of interest. The fact that diversity and craftmanship need more time to be predicted (longer interaction lengths) can be explained by the fact that a more diverse website is likely to increase interest and attention of the users [2, 3], making them look for details and spend more time looking at it. Similarly, for craftmanship, users may need to look at the website more carefully, thus

need more time, in order to make judgements if a website is well built and of high quality.

This study contributes both theoretically and practically. We contribute to the need of automatizing the process of capturing and explaining users' perceptions towards visual aesthetics. The use of objective data from eye-tracking combined with questionnaires can remove bias and offer more robust results. Drawing on recent discussion in HCI on how machine learning and AI can improve the field [8] as well as to be used to improve image [9] and website aesthetics [10], we offer empirical evidence on how the process of predicting users' preferences can be automated. The findings pave the way to design new systems that will take into account users' gaze behavior in an unobtrusive manner and will be able inform researchers and designers about their perceptions of visual aesthetics. An AI application or an intelligent agent may compute in real time the probability for a user to find a website high or low appealing, which in turn may be used as feedback to improve design. Furthermore, building on previous studies showing that users' gaze is related both with their perceptions and cognitive processing [18, 19], we provide empirical evidence that gaze behavior is linked to perceptions towards the same stimuli.

Managers and websites designers can benefit from our findings and use them to better understand their customers with a focus on their perceptions of visual appeal. As technology evolves, eye-tracking devices are becoming cheaper and accessible to larger audiences. At the same time, machine learning and AI have evolved at a level where we see their applications in our daily lives, and with automated machine learning (AutoML) practitioners can more easily develop such new applications. For example website retailers can use eye tracking or even users' own mobiles phones to evaluate websites [35]. By knowing how quickly we can predict users' ratings, a developer can design an application that will request to track users' eye behavior only for the first 15 or 20 seconds of using the application. This could serve as a middle ground, where the company will get important data to improve their application and users' data will be recorded at a minimum level, making it more likely for users to accept and share their data. This can be combined with knowledge on gaze behavior, as they can show accurately which parts of a website were more or less important (e.g., high fixation on specific areas or high saccade velocity between two areas) and how that relates to their ratings regarding aesthetics. Thus, if we can predict high ratings on simplicity then we can automatically identify which areas created this perception to the user.

The present study has some limitations. First, we measure visual aesthetics through questionnaires, thus taking a subjective approach [2]. Future studies, may combine subjective and objective methods, such as using screen design factors [1], along with eye-tracking techniques to acquire a deeper understanding of users' judgments of visual appeal. Furthermore, although the sample of the study is relatively small, the low-level data that we capture are able to offer useful insight, allowing us to quantify aesthetic qualities using eye-tracking measures [22]. Finally, to better understand the users and their online experience, complexity theory may be combined with the novel fuzzy-set qualitative comparative analysis (fsQCA) as it can offer deeper understanding of the sample by identifying asymmetric relations inside a dataset [36, 37].

10

Appendix

Items to measure visual aesthetics

Simplicity
The layout appears too dense (r).
The layout is easy to grasp.
Everything goes together on this site.
The site appears patchy (r).
The layout appears well structured.
Diversity
The layout is pleasantly varied.
The layout is inventive.
The design appears uninspired (r).
The layout appears dynamic.
The design is uninteresting (r).
Colorfulness
The color composition is attractive.
The colors do not match (r).
The choice of colors is botched (r).
The colors are appealing.
Craftsmanship
The layout appears professionally designed.
The layout is not up-to-date (r).
The site is designed with care.
The design of the site lacks a concept (r).

References

1. Seckler, M., Opwis, K., Tuch, A.N.: Linking objective design factors with subjective aesthetics: An experimental study on how structure and color of websites affect the facets of users' visual aesthetic perception. Computers in Human Behavior 49, 375-389 (2015)
2. Moshagen, M., Thielsch, M.T.: Facets of visual aesthetics. International Journal of Human-Computer Studies 68, 689-709 (2010)
3. Tuch, A.N., Bargas-Avila, J.A., Opwis, K.: Symmetry and aesthetics in website design: It'sa man's business. Computers in Human Behavior 26, 1831-1837 (2010)
4. Lavie, T., Tractinsky, N.: Assessing dimensions of perceived visual aesthetics of web sites. International journal of human-computer studies 60, 269-298 (2004)
5. Hartmann, J., De Angeli, A., Sutcliffe, A.: Framing the user experience: information biases on website quality judgement. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 855-864. ACM, (Year)
6. Reimann, M., Zaichkowsky, J., Neuhaus, C., Bender, T., Weber, B.: Aesthetic package design: A behavioral, neural, and psychological investigation. Journal of Consumer Psychology 20, 431-441 (2010)

7. Bhandari, U., Neben, T., Chang, K., Chua, W.Y.: Effects of interface design factors on affective responses and quality evaluations in mobile applications. Computers in Human Behavior 72, 525-534 (2017)

8. Inkpen, K., Chancellor, S., De Choudhury, M., Veale, M., Baumer, E.P.: Where is the Human?: Bridging the Gap Between AI and HCI. In: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, pp. W09. ACM, (Year)

9. Lu, X., Lin, Z., Jin, H., Yang, J., Wang, J.Z.: Rating image aesthetics using deep learning. IEEE Transactions on Multimedia 17, 2021-2034 (2015)

10. Chen, A., Nah, F.F.-H., Chen, L.: Assessing Classical and Expressive Aesthetics of Web Pages using Machine Learning. MWAIS 2018 Proceedings, Saint Louis, Missouri, US (2018)

11. Djamasbi, S., Siegel, M., Tullis, T.: Generation Y, web design, and eye tracking. International journal of human-computer studies 68, 307-323 (2010)

12. Djamasbi, S., Siegel, M., Tullis, T., Dai, R.: Efficiency, trust, and visual appeal: Usability testing through eye tracking. In: System Sciences (HICSS), 2010 43rd Hawaii International Conference on, pp. 1-10. IEEE, (Year)

13. Hwang, Y.M., Lee, K.C.: Using an Eye-Tracking Approach to Explore Gender Differences in Visual Attention and Shopping Attitudes in an Online Shopping Environment. International Journal of Human–Computer Interaction 1-10 (2017)

14. Pappas, I, O, Sharma, K., Mikalef, P., Giannakos, M.N.: Visual Aesthetics of E-Commerce Websites: An Eye-Tracking Approach. Proceedings of the 51st Hawaii International Conference on System Sciences, Big Island, Hawaii (2018)

15. Pappas, I.O., Sharma, K., Mikalef, P., Giannakos, M.N.: A Comparison of Gaze Behavior of Experts and Novices to Explain Website Visual Appeal. In: PACIS, pp. 105. (Year)

16. Lindgaard, G., Fernandes, G., Dudek, C., Brown, J.: Attention web designers: You have 50 milliseconds to make a good first impression! Behaviour & information technology 25, 115-126 (2006)

17. Bhandari, U., Chang, K., Neben, T.: Understanding the impact of perceived visual aesthetics on user evaluations: An emotional perspective. Information & Management 56, 85-93 (2019)

18. Just, M.A., Carpenter, P.A.: A theory of reading: From eye fixations to comprehension. PsychologR 87, 329 (1980)

19. Arnheim, R.: Visual dynamics. Am. Sci. 76, 585-591 (1988)

20. Livio, M.: The golden ratio: The story of phi, the world's most astonishing number. Broadway Books (2008)

21. Wang, Q., Yang, S., Liu, M., Cao, Z., Ma, Q.: An eye-tracking study of website complexity from cognitive load perspective. Decision support systems 62, 1-10 (2014)

22. Khalighy, S., Green, G., Scheepers, C., Whittet, C.: Quantifying the qualities of aesthetics in product design using eye-tracking technology. Int. J. Ind. Ergonomics 49, 31-43 (2015)

23. Mikalef, P., Sharma, K., Pappas, I.O., Giannakos, M.N.: Online Reviews or Marketer Information? An Eye-Tracking Study on Social Commerce Consumers. In: Conference on e-Business, e-Services and e-Society, pp. 388-399. Springer, (Year)

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

25. Prieto, L.P., Sharma, K., Kidzinski, Ł., Rodríguez-Triana, M.J., Dillenbourg, P.: Multimodal teaching analytics: Automated extraction of orchestration graphs from wearable sensor data. J. Comput. Assisted Learn. 34, 193-203 (2018)

26. Reichle, E.D., Warren, T., McConnell, K.: Using EZ Reader to model the effects of higher level language processing on eye movements during reading. Psychonomic bulletin & review 16, 1-21 (2009)

27. Jaarsma, T., Jarodzka, H., Nap, M., van Merrienboer, J.J., Boshuizen, H.P.: Expertise under the microscope: processing histopathological slides. Medical education 48, 292-300 (2014)

28. Abernethy, B., Russell, D.G.: Expert-novice differences in an applied selective attention task. Journal of Sport and Exercise Psychology 9, 326-345 (1987)

29. Krischer, C., Zangemeister, W.H.: Scanpaths in reading and picture viewing: Computer-assisted optimization of display conditions. Comput. Biol. Med. 37, 947-956 (2007)

30. Zangemeister, W.H., Liman, T.: Foveal versus parafoveal scanpaths of visual imagery in virtual hemianopic subjects. Comput. Biol. Med. 37, 975-982 (2007)

31. Liao, K., Kumar, A.N., Han, Y.H., Grammer, V.A., Gedeon, B.T., Leigh, R.J.: Comparison of velocity waveforms of eye and head saccades. Annals of the New York Academy of Sciences 1039, 477-479 (2005)

32. Russo, M., Thomas, M., Thorne, D., Sing, H., Redmond, D., Rowland, L., Johnson, D., Hall, S., Krichmar, J., Balkin, T.: Oculomotor impairment during chronic partial sleep deprivation. Clinical Neurophysiology 114, 723-736 (2003)

33. Phillips, M.H., Edelman, J.A.: The dependence of visual scanning performance on search direction and difficulty. Vision Res. 48, 2184-2192 (2008)

34. Vuori, T., Olkkonen, M., Pölönen, M., Siren, A., Häkkinen, J.: Can eye movements be quantitatively applied to image quality studies? In: Proceedings of the third Nordic conference on Human-computer interaction, pp. 335-338. ACM, (Year)

35. Zhu, A., Wei, Q., Hu, Y., Zhang, Z., Cheng, S.: MobiET: A New Approach to Eye Tracking for Mobile Device. In: Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers, pp. 862-869. ACM, (Year)

36. Pappas, I.O., Giannakos, M.N., Sampson, D.G.: Fuzzy set analysis as a means to understand users of 21st-century learning systems: The case of mobile learning and reflections on learning analytics research. Computers in Human Behavior 92, 646-659 (2019)

37. Pappas, I.O.: User experience in personalized online shopping: a fuzzy-set analysis. European Journal of Marketing 52, 1679-1703 (2018)