

Visual Aesthetics of E-Commerce Websites: An Eye-Tracking Approach

Ilias O. Pappas
Norwegian University
of Science and
Technology
ilpappas@ntnu.no

Kshitij Sharma
École Polytechnique
Fédérale de Lausanne
& Norwegian
University of Science
and Technology
kshitij.sharma@epfl.ch

Patrick Mikalef
Norwegian University
of Science and
Technology
patrick.mikalef@ntnu.no

Michail N. Giannakos
Norwegian University
of Science and
Technology
michailg@ntnu.no

Abstract

This study adopts four facets of visual aesthetics (i.e., simplicity, diversity, colorfulness, craftsmanship) to explain how they relate with users' gaze patterns, based on how much they fixate on certain points, as well as how fast and how much distance their eyes cover. On a sample of 23 experienced users in online shopping, we collect eye-tracking data while looking at high, neutral, and low appealing websites, and then register their perceptions on visual aesthetics towards those websites. Findings show different patterns of gaze behavior related with users' perceptions on visual aesthetics. Short fixations with high saccade show high simplicity, while high fixation variance and high backtrack shows high diversity. Short fixations with high backtrack show high colorfulness. Low saccade velocity with high skewness shows high craftsmanship. We contribute towards the need of automatizing the process of understanding users' perceptions of visual aesthetics, as we might be able to predict the user behavior in real time in the future.

1. Introduction

As e-commerce market grows, existing retailers always consider improving their websites to attract more customers, and new retailers aim to develop websites that are not only effective and visually appealing, but also different from existing ones. Multiple studies in Human Computer Interaction (HCI) have examined website design characteristics including their effectiveness and visual aesthetics [1, 2, 3, 4]. By visiting the top-100 websites in online shopping, in terms of popularity (based on ratings from alexa.com), it is evident that online retailers choose different ways to design their homepages, raising the need for further investigation on what customers consider to be visually appealing to them.

The role of website aesthetics in online shopping websites has been mainly examined by taking subjective approach, which is based on questionnaires [2], or a more objective approach, which is based on screen design factors and layout elements linked with perceptions of visual aesthetics [5]. Furthermore, recent studies have taken these two approaches together to offer a deeper insight on what affects users' visual aesthetic perceptions [1]. Nonetheless, data received from the users remain to be subjective, increasing the possibility of bias, which can be reduced by utilizing neurophysiological methods, such as fMRI and eye-tracking [6, 7].

Recent studies employ eye-tracking techniques to capture users' eye movement and gaze to identify specific areas of interest when looking at a website [8, 9, 10]. However, since most of these studies present heat maps to show gaze behavior it is less clear how the users looked at the website. By reading a heap map we can see where users focused, however we do not know how their eyes moved the whole time (e.g., speed, change in direction, switch between two spots), which can give more insight on how they looked at the website and how appealing they perceive it to be.

Consumers' visual attention is related with the cognitive processing of the information displayed on a website [10]. Online retailers design their websites to capture attention, but also to be appealing in the eyes, while containing all the necessary information one is looking for. However, it remains unclear how users look at a website, and specifically what are their eye movements when evaluating a homepage for the first time. Thus, we propose the following research question: *RQ: How does gaze behavior relate to the perception of visual aesthetics in websites?*

To answer our research question, we explore how users perceive visual websites' aesthetics when looking at e-commerce websites. We conducted an experiment in which we first categorized websites into high, neutral, and low appealing with the help of experts in design and HCI, and then measured the eye

tracking behavior of users' while looking at these websites, along with their perceptions towards different facets of visual aesthetics (i.e., simplicity, diversity, colorfulness, craftsmanship). We contribute by combining objective gaze behavior with subjective aesthetic perceptions, offering a better understanding on how users' look at various websites. By doing so, we provide evidence of how subjective perceptions are formed based on objective eye-tracking data. Also, we contribute to the need of automatizing the process of understanding users' perceptions of visual aesthetics without the need to ask them, thus reducing the subjective bias of questionnaires, since gaze behavior may be used to predict if a user likes or not a website.

2. Background

2.1. Visual Aesthetics

Recent studies in HCI identify various attributes and characteristics that influence users' perceptions towards the aesthetic evaluation of websites [5, 2, 1, 3]. Design approaches link screen design factors and layout elements to users' perception of visual aesthetics [3]. Visual aesthetics have been measured using questionnaires, thus taking a subjective take on the matter. Visual aesthetics have been categorized into classical and expressive aesthetics [4]. Classic aesthetics refer to clear, organized, and pleasant websites, while expressive aesthetics refer to creative, colorful, and original websites.

To better describe the domain of visual aesthetics, a new measure was developed, based on four interrelated facets of perceived visual aesthetics, offering a more complete depiction of aesthetics [2]. The four facets are *simplicity*, *diversity*, *colorfulness*, and *craftsmanship*. Simplicity refers to how a website is designed in terms of unity, clarity, orderliness and balance, and is related to the classical aesthetics. Diversity refers to the website being dynamic, novel and creative, and is related to the expressive aesthetics. Extending the work of Lavie and Tractinsky [4], colorfulness and craftsmanship were added to evaluate the unique effect of colors, and the skilful and coherent integration of all design dimensions, respectively [2, 1]. Here, we adopt the measures created by Moshagen and Thielsch [2] as they are more detailed and include more aspects of aesthetics, thus offering a deeper understanding of perceived visual aesthetics.

2.2. Eye-tracking

To measure the visual appeal of websites recent studies have employed eye-tracking techniques to

capture the eye movement of the users when looking at specific websites [8], which may offer a better understanding of user behavior when visiting e-commerce websites. Indeed, users' perceptions towards visual aesthetic appeal can be predicted by employing functions which are based on specific measurable qualities of a stimulus. Building on this notion, various website design characteristics as well as visual attention and attitude have been explained by using users' gaze behavior [8, 10].

Recent studies link website complexity with gaze behavior, and show that there is no significant difference among high, medium, and low complex websites for fixation duration [11]. This suggests that users' gaze is at similar levels regardless of how complex (or simple) a website is, when they perform simple tasks. Interestingly, for complex tasks fixation duration is higher for sites with medium complexity, compared with high and low [11], indicating that more data from the gaze behavior are needed, such as the speed that the eye moves or the times that it moves back and forth between two points, which offer more information on the user looks at the website.

Visual appeal depends on various characteristics, like size of pictures, search functionality, text quantity, and the fixation duration has been used to measure where users focus when looking at a website [8]. Although such factors are important for a website as they attract the attention of the users, they offer limited insight on users' perceptions towards simplicity, diversity, colorfulness, and craftsmanship, which are deemed important in measuring website visual aesthetics [2, 1, 4]. Thus, although it is highly relevant for website design, there has been little research in linking the aforementioned factors with gaze behavior, which can provide empirical support on how to improve website design or on how to deduce users' perceptions of visual aesthetics without asking them, based on the way they look at the website.

2.3. Visual aesthetics and Eye-tracking

This study offers a quantitative point-of-view about visual aesthetics using eye-tracking data. The aesthetical preferences of a person can be predicted with the derivatives of the stimulus properties [8, 12, 13], or in other words "what you perceive is what you see". Also, based on the eye-mind hypothesis [14], what a person looks at indicates ones cognitive process, or in other words "what you see is what you process". Indeed, eye-tracking research links gaze with the cognitive load theory [15, 11], supporting the eye-mind hypothesis [14]. Thus, what one perceives and what one processes is linked with his or her gaze. We use gaze variables as a proxy for the visual

aesthetics perception, to examine how looking at websites explains perceptions of visual appeal.

A typical set of eye-tracking measures for cognitive load [16] includes pupil diameter, fixation duration and its variance, and saccadic length and velocity. Here, we avoid pupil diameter, susceptible to many variations caused by external factors [17], and add two additional variables (i.e., backtracks, and saccade velocity skewness) that have been found to correlate with the anticipation patterns [18, 19]. Finally, we examine the anticipation patterns because anticipation plays a key role in visual perception [20].

3. Experimental Material

Following previous studies in the area we use the homepages of popular online shopping websites to examine visual appeal [8]. To choose the websites, the top 100 websites, based on popularity, were retrieved from Alexa.com, which offers global traffic estimates in various categories. In this case, the category of shopping was chosen. Next, 50 websites were chosen randomly, and 10 experts in HCI and website design were asked to evaluate on a 1-7 scale how appealing they find them. Then, the websites were ranked, and three categories were created based on the median; high, neutral, and low visually appealing websites. Finally, 9 websites were chosen, 3 for each category. Figure 1 shows examples from each category.

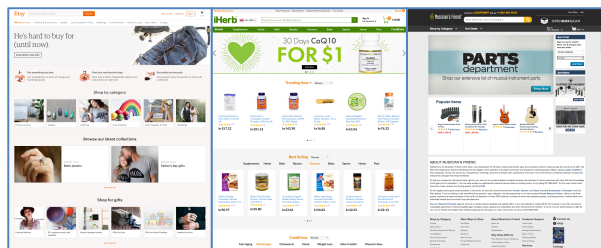


Figure 1 High, Neutral, Low Appealing e-commerce websites (from left to right)

By observing these examples, it is clear that a main difference among the three categories is the ratio of pictures and text in the homepage, that is a key design characteristic in visually appealing websites [8, 2]. In detail, the high appealing website has the least text accompanying the pictures, the neutral has more text and the low appealing homepage has less pictures than text. Previous studies show that the first impression of a webpage is formed almost immediately [21], thus investigating users' gaze and focus in the first seconds that they look at a webpage is relevant and may prove useful in recognizing how their behavior is related with key factors of visual aesthetics.

4. Method

4.1. Participants and procedure

There were 23 participants in the experiment; 10 females and 13 males. The average age was 27.5 years (S.D. 7.15 years). All were experienced with online shopping. Participants were recruited through emails and social media. The study was conducted at a dedicated lab space of a large university in Norway. The lab was used by 5 users at a time.

First the participants reviewed and signed an informed consent document about this study. Basic information was collected, including age, gender, and internet experience, and the researchers explained the objective of the study. The eye tracker was calibrated to the participant, a brief procedure in which the participant watched a dot move to the four corners and the center of the screen. The calibration process for each user took about fifteen seconds. Participants were given instructions to view each web page for a minimum of ten seconds, and then provide appeal ratings for the four facets of visual aesthetics, on a 7-point scale. To control for which websites were viewed and when, the websites were presented in a specific order (high, neutral, low appeal), making sure that all participants viewed them in the same order. Participants were told the visual appeal rating was not to be based on the content of the page. Also, as having to perform a task influences the way users look at the website and where they focus, we did not give specific tasks [11], instead we asked them to look at the homepage freely. Participants controlled the mouse to scroll and to choose when to continue to the next page.

4.2. Eye-tracking device

Four SMI RED 250 and one TOBII mobile eye-tracker were used. A sampling rate of 60hz is sufficient for usability studies [22]. The mobile eye-trackers are worn as glasses. The frames of eye-tracking glasses contain two cameras capturing each eye of the participants; and an array of infrared LEDs which are reflected by the cornea. Another camera captures the field of view of the participants. Based on these the software calculates the position of the gaze on the field of view video. We rectify for head movements with fiducial markers placed on the laptop screen.

4.3. Measurements and analysis

Regarding aesthetics and visual appeal we examine simplicity, diversity, colorfulness, and craftsmanship [2]. Perceived visual aesthetics are measured by a 7-

point Likert scale, and then three categories are created for the low/medium/high scores, by computing the 33% and 67% percentile. The three categories are chosen based on categories created from experts' ratings. Table 1 shows cut-off points for each type of website (i.e., high, neutral, low appeal), and theoretical limits (min, max values) of each factor. The limits are computed based on 7-point scale and the fact that some items are reverse coded. Items used to measure the constructs are presented in the Appendix.

Table 1. Thresholds for visual aesthetics

Visual aesthetics (min, max)	High Appeal	Neutral Appeal	Low Appeal
Simplicity (-11, 19)	7, 13	5, 9	2, 5
Diversity (-11, 19)	7, 11	1, 9	-3, 3
Color (-12, 12)	3, 8	0, 6	-4, 1
Craft (-12, 12)	4, 8	1, 6	-5, 1

From the eye-tracking data we extracted the following variables.

Fixation Duration. This is the average period of time during which the gaze is relatively stable over a relatively longer period of time, measured in milliseconds. This is often correlated to the careful investigation of the stimulus. Just and Carpenter [14] proposed an eye-mind hypothesis, stating that there is no substantial lag between what we fixate on and what we process. To a large extent, the past research had supported this hypothesis in different tasks, for example, in reading and scene perception [23, 24, 25]. According to Unema, et al. [26] shorter fixations are categorized as global processing, while longer fixations as local processing. For having a global view of the stimulus participants used, what is called, "move and fixate" behavior which incurs in shorter fixations.

Saccadic Amplitude. This is the distance travelled by the eyes in a relatively short period of time while jumping from one place to another. Longer saccadic amplitudes reflect overview scans processing while shorter saccadic amplitudes reflect careful scans [27].

Saccadic Velocity. This is the speed that eyes jump from one place to another. Anticipations are often accompanied by lower saccadic velocity [18, 28].

Saccadic Velocity Skewness. This is the skewness of the velocity histogram of every participant. The higher the skewness, the higher the proportion of anticipatory saccades in eye-movements [18].

Ratio backtracks. A backtrack is a combination of two saccades when the second one goes in the opposite direction of the first one. Backtracks are indicative of

revisiting previously visited parts of the stimulus [19]. We computed the ratio of the backtracks as the proportion of the total number of saccades. Table 2 shows descriptive statistics for eye-tracking variables.

Table 2. Descriptives of eye-tracking variables

Eye-tracking variables	Mean	SD
Fixation duration mean	150.84	34.82
Fixation duration SD	136.33	52.01
Saccade amplitude	4.68	4.17
Saccade Velocity	40.09	21.85
Saccade velocity skewness	7.14	7.31
Ratio backtrack	0.27	0.07

As the outcome has a limited number of possible values, logistic regression is chosen for data analysis. Also, multiple linear regression was performed to compare findings. Both techniques have the same outcome. Overall appeal ratings from experts are the independent variable and the end-users' ratings are the dependent variable. We include experts' ratings in the model along with eye-tracking data, for two reasons; to give quantitative measures for aesthetic perception, and to compare end-users' and experts' ratings.

5. Results

Tables 3-6 show users' eye tracking behavior when looking at websites, for all four facets of visual aesthetics and for all three categories of websites, as well as the differences between low, medium, and high ratings in user responses. As we need to capture gaze patterns from different viewpoints to examine the relation between aesthetics and eye-movements, the findings present only the best models for each visual aesthetic. Indeed, the models presented in Tables 3-6 are the best fitting models based on Akaike Information Criterion (AIC)¹ values of different models. Thus, we see different set of gaze variables explaining different facets of visual aesthetics. Also, to examine if including Appeal variable in the analysis leads to a better model we tested the full model, and the model with and without appeal. Findings show that the full model has the worst fit among the three (highest AIC) in explaining the outcome. Also, the model with appeal has the lowest AIC, being statistically different from the model without appeal.

Table 3 reads as follows: An increase in fixation duration mean is associated with a decrease in the log odds² of giving a medium vs low rating with 0.05, and giving a high vs low rating with 0.07. Next, an increase

the model. Thus, the model incurring in the lowest information loss will be the best model and will have the lowest AIC value.

²Odds are a measure for the strength of association between presence/absence of event A and presence/absence of event B.

¹AIC is a quality measure of a statistical model. It offers a way to select the most suitable model (better fit) for a dependent variable. It represents the information loss while using certain variables in

in saccade amplitude is associated with the increase in the log odds of giving a medium vs low rating with 1.06, and giving a high vs low rating with 1.15. Regarding appeal, the increase (or decrease) in the log odds refers to users giving different ratings of high/medium/low for all types of websites coded using experts' ratings. The log odds of giving a medium vs low rating will increase by 8.28 if a website has neutral appeal. The log odds of giving a high vs low rating will increase by 8.01 if a website has high appeal. The log odds of giving a high vs low rating will increase by 9.82 if a website has neutral appeal. The log odds of giving a high vs low rating will increase by 7.76 if a website has low appeal. Table 3 shows that higher simplicity ratings correspond to shorter fixations and longer saccades, as shorter fixations and longer saccades depict global processing. For all types of appeal, the increase in probability of someone giving high vs low rating is highest in neutral appeal and lowest in low appeal. The intercept of the model was forced to be zero, thus all ratings are equiprobable.

Table 3. Logistic model results to explain simplicity with gaze data

Simplicity	Coefficients		z-score	
	Medium /Low	High /Low	Medium /Low	High /Low
Fixation duration mean	-0.05	-0.07	-2.64***	-3.17***
Saccade amplitude	1.06	1.15	2.00*	2.18*
High Appeal	6.22	8.01	1.71*	2.22*
Neutral Appeal	8.28	9.82	1.96*	2.34*
Low Appeal	4.83	7.76	1.28	2.09*

Fixation is measured in msec. Saccade in degrees. The intercept was forced to be zero for better interpretability.
 *** p < .001, ** p < .01, * p < .05

Figure 4 shows how the probability of giving a rating about simplicity changes with fixation duration. The probability of giving high scores in terms of simplicity to a website reduces as fixation duration increases. On the other hand, the probability of giving low scores to a website increases as fixation duration increases. This is observed for all types websites (i.e., high, low, neutral appeal). Thus, the more the users fixate while looking at a webpage the more details they view, and the less simple they find the website.

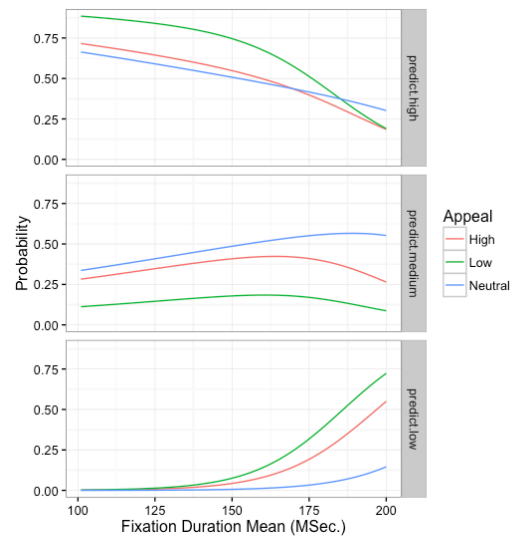


Figure 2. Evolution of users' simplicity ratings (high/medium/low), based on fixation duration mean and experts' ratings.

Table 4 shows users' gaze behavior for diversity. In detail, higher diversity ratings are coincided with shorter fixation means and high fixation duration variance. This suggests that there might be periods of long fixations and short fixations indicating a mixture of global and local processing. Also, for all types of appeal, the decrease in probability of someone giving a high rating vs a low rating is highest in low appeal and the lowest in neutral appeal.

Table 4. Logistic model results to explain diversity with gaze data

Diversity	Coefficients		z-score	
	Medium /Low	High /Low	Medium /Low	High /Low
Fixation duration mean	-0.037	-0.05	-2.16*	-2.94**
Fixation duration SD	0.031	0.047	2.00*	2.85**
Ratio backtracks	22.4	33.6	4.37***	7.99***
High Appeal	-4.24	-6.25	-1.35	-1.96*
Neutral Appeal	-3.85	-6.65	-1.14	-1.94
Low Appeal	-4.32	-6.09	-1.50	-2.10*

Fixation is measured in msec. Saccade in degrees. The intercept was forced to be zero for better interpretability.
 *** p < .001, ** p < .01, * p < .05

Figure 5 shows how the probability of giving a rating for diversity changes with fixation duration. Like for simplicity, the probability of giving a high or low score to a website decreases and increases, respectively, as fixation duration increases. Yet, for medium ratings a decrease is observed, unlike to simplicity. The closer the users look at a website the more likely is for perceptions on diversity to drop

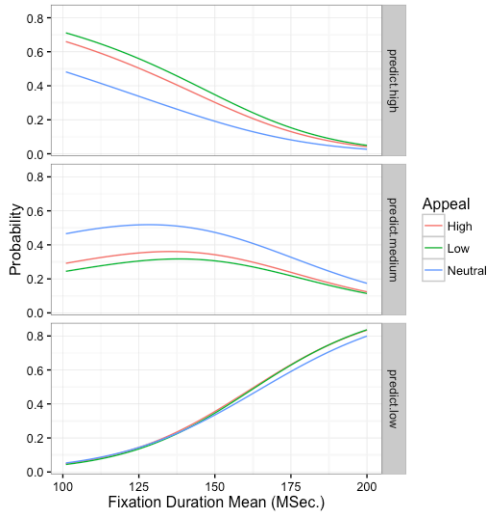


Figure 3. Evolution of users' diversity ratings (high/medium/low), based on fixation duration mean and experts' ratings.

Table 5 shows users' gaze behavior for colorfulness. Higher colorfulness ratings appear with short fixations and high backtracking ratio.

Table 5. Logistic model results to explain colorfulness with gaze data

Colorfulness	Coefficients		z-score	
	Medium /Low	High /Low	Medium /Low	High /Low
Fixation duration mean	-0.03	-0.03	-2.51**	-2.54**
Ratio backtracks	10.55	15.38	1.79	2.53**
High Appeal	2.14	0.86	0.85	0.35
Neutral Appeal	3.00	1.59	1.09	0.59
Low Appeal	1.73	0.67	0.70	0.28

Fixation is measured in msec. Saccade in degrees. The intercept was forced to be zero for better interpretability.
 *** p < .001, ** p < .01, * p < .05

Table 5 suggests that participants spent time in comparing different parts of the website resulting in back and forth saccades and short fixations. Yet, we notice no effect of appeal on the ratings concerning colorfulness. The findings show that color did not make a difference on how the users rated each type of website.

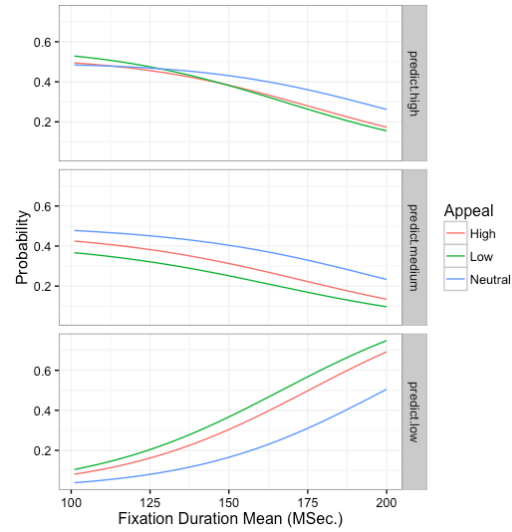


Figure 4. Evolution of users' colorfulness ratings (high/medium/low), based on fixation duration mean and experts' ratings.

Figure 6 shows that for colorfulness the probability of giving a high or low score to a website decreases and increases, respectively, as fixation duration increases. Also, for medium ratings a decrease is observed as well. However, the graph shows that the probability to give a lower rating is much higher and increases faster, as fixation duration increases, compared to medium and high ratings. This suggests it is easier to identify and categorize colors that are less appealing.

Finally, Table 6 presents users' gaze behavior for craftsmanship. Higher craftsmanship is related to lower velocity and a skewer velocity distribution, which is often correlated with the anticipatory gaze. Also, for all types of appeal, the increase in probability of someone giving a high rating vs a low rating is highest in neutral appeal and the lowest in high appeal.

Since the skewness of saccade velocity indicates towards anticipatory gaze from the user, we show in Figure 7 how the increase in skewness is related with the probability of giving high, medium, or low scores on the websites. In detail, for low and neutral appealing websites, an increase in skewness reduces the probability of giving low ratings up to almost zero. This suggests that when users know what they expect

to see, indicating a familiarity with the website, the chance to perceive the website as low appealing are reduced. On the other hand, the probability to give medium or high scores to all types of websites increases, as the saccade velocity skewness increases. Thus, anticipatory eye movements suggest that the users will give higher ratings to the websites.

Table 6. Logistic model results to explain craftsmanship with gaze data.

Craftsmanship	Coefficients		z-score	
	Medium /Low	High /Low	Medium /Low	High /Low
Saccade velocity	-0.05	-0.08	-2.11*	-3.16**
Saccade velocity skewness	0.18	0.19	2.13*	2.22*
High Appeal	0.74	2.26	0.62	2.04*
Neutral Appeal	2.03	3.20	1.31	2.02*
Low Appeal	0.71	2.28	0.58	1.93

The intercept was forced to be at zero for better interpretability. *** $p < .001$, ** $p < .01$, * $p < .05$

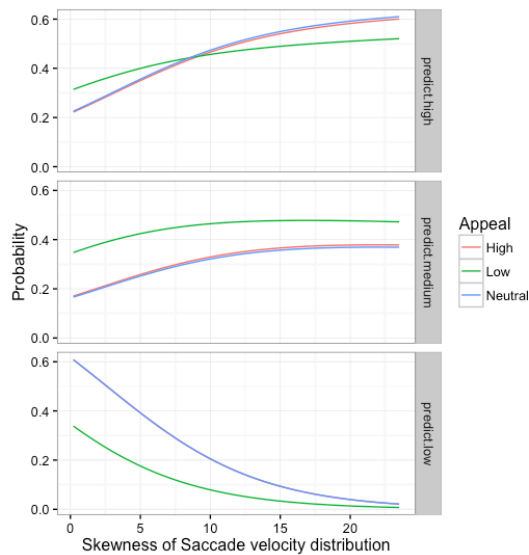


Figure 5. Evolution of users' craftsmanship ratings (high/medium/low), based on fixation duration mean and experts' ratings.

6. Discussion

Following the need of constantly improving the visual appearance of e-commerce websites, we offer a

deeper explanation of how consumers look at such websites and how gaze relates with perceived visual aesthetics. Building on two basic notions, that is “what you perceive is what you see” [12, 13] and “what you see is what you process” [14], we posit that gaze behavior can predict users’ visual aesthetic perceptions. Indeed, the findings indicate that with the advent in our understanding of the relations between users’ ratings (aesthetics) and gaze patterns, we can explain consumers’ cognitive load while visiting websites and predict their behavior in real time.

We adopt four central characteristics of website design (i.e., simplicity, diversity, colorfulness, and craftsmanship) and link users’ ratings for each one with eye-movement, for high, neutral, and low appealing e-commerce websites. Analyzing data at a low level of abstraction (fixation and saccade based variables), allows us to explain behavior with less amount of data. We track eye-movement (i.e., fixation duration, saccadic amplitude, saccadic velocity, saccadic velocity skewness, ratio backtracks) to explain users’ ratings on visual aesthetics of websites. Table 7 presents a summary of the findings.

Table 7. Summary of findings

Visual aesthetic	Eye movement	Main finding / Rationale
Simplicity	Short fixation duration means & high saccade amplitude	-High simplicity -Getting an overview of the website (global processing) -No focus on details on the first visit
Diversity	High fixation duration variance (SD) & high backtrack	-High diversity -Focus on the details -Increased interest & cognitive load
Colorfulness	Short fixation duration means & high backtrack	-High colourfulness -Compare colors of website & products -Expect high appealing websites to match - look in order.
Craftsmanship	Low saccade velocity with high skewness	-High craftsmanship -Seeking pre-assumed characteristic to evaluate craftsmanship -Anticipatory gaze -Eye-movements are driven by a preexisting mental model of user

The findings show that when users have short fixations and move their eyes fast from point to point, they tend to get an overview of the website (i.e., global processing) and perceive it to be simpler, suggesting

that they do not focus on the details when they first visit the website. Indeed, since a simple design can be processed more fluently and effortlessly, it can be explained why users take an overview of the website instead of fixing their gaze on more details. Further, when users show a combination of both global and local processing (i.e., short fixation duration means with high variance) they perceive the website to be more diverse. Different from simplicity, diversity seems to be interpreted by users as giving focus to more details, hence the combination between global and local processing. This is explained by the fact that the existence of diversity increases interest and tension of the user [2, 3], thus making the user to watch more closely the website and identify more details.

Next, when users fixate for short durations while at the same time they move their gaze often back and forth between two points, they are likely to perceive the website as more colorful. Since most e-commerce websites have multiple products in their homepages, which change often, the colors of the homepage change often as well. Colorfulness is based on perceptions that arise from selection, placement, and combination of the colors [2]. Thus, users compare the colors in the website and the colors of the products, which attract their attention, and expect them to match and look in order on high visually appealing websites. This comparison can be easier when existing colors match poorly or improperly, leading the user to form low perceptions of visual appeal faster (i.e., with low fixation duration mean).

Moreover, looking at a website more slowly combined with an anticipatory gaze the users are likely to find the website to be well built, as a skillful and coherent integration of all design characteristics leads to perceptions of high craftsmanship [2]. Furthermore, this eye behavior suggests that users show anticipatory gaze, which means that they were seeking some pre-assumed characteristic in the webpages to evaluate the craftsmanship. This shows a typical gaze pattern where the eye-movements are driven by the preexisting mental model of the participants.

Our findings indicate that users tend to give higher scores on visual aesthetics than it was expected. Since the websites have been categorized following HCI and web design experts' recommendations, we expected low ratings for low appealing websites, medium for neutral websites, and high for high appealing websites. Of course, users have various opinions and tastes, which could partially explain the above-mentioned tendency of giving higher scores. However, this is a concern for the usability research, since higher than expected ratings might mislead the design. Using the proposed model, we can correct for the subjectivity

from different users and offer a more objective view on the usability (aesthetics in the present contribution).

The findings also suggest that the more users' focus on the website, the more likely it gets to identify less appealing details, which lead to lower scores on visual aesthetics. Additionally, when users' stop comparing the different characteristics of the website, it could be an indication that the website is perceived to be low appealing. Previous studies have shown that first impressions are shaped in the first seconds that users look at a website, thus any aspect grabbing their attention first will have a great influence on their perceptions towards visual appeal [21]. However, our findings show that the longer users fixate, the probability to rate the website with a specific score changes, indicating a possible change on their first impressions. For example, the probability to give a high score to a high appealing website drops, while the probability to give a low score to a low appealing website increases, as fixation duration increases.

6.1. Implications

This study has both theoretical and practical implications. First, it shows that users' gaze patterns of a stimuli are related with their perceptions towards the same stimuli, extending previous theoretical work that shows users' gaze is linked both with their perceptions and cognitive processing [12, 14]. Second, it complements and extends previous research in the area of visual aesthetics, by examining four critical facets of visual aesthetics, and by employing eye-tracking method to explain how users' look at websites and how they rate them in terms of visual appeal. Third, we add to current studies highlighting the importance of eye-tracking in user behavior, and go beyond the qualitative studies that use fixation duration to create heat maps [8, 11], by examining more measures, like saccadic amplitude, saccadic velocity, and ratio backtracks. Such measures offer a quantitative connection among visual aesthetics and gaze data, for a deeper insight on how users' look at a website, improving our understanding of user behavior. Finally, we add to the need to automatize the process of understanding users' perceptions of visual aesthetics. Since questionnaires are subjective, being able to capture objective data like the eye-movements, may prove useful for both research and practice.

After analysing gaze behaviour we have a concrete model of how eye-movements can predict users' ratings. Feeding this model into an intelligent agent will compute in real-time the probability for a user to give high/medium/low rating. This can help to predict users' ratings and give proper feedback about their misconceptions. For example, in a website with high

craftsmanship, when the agent predicts that users will give a low rating, it could ask them to describe their views of what they consider as high craftsmanship and then guide them to the specific areas of the webpage to compare their beliefs with the actual content.

Practitioners and designers of e-commerce websites may benefit from our findings and use them to improve the understanding their customers and specifically their perceptions of visual appeal. In effect, we provide empirical support that gaze behavior is related with perceptions of visual aesthetics. Online retailers, especially new ones that are unknown to most of the consumers, should strive to improve their homepages as they play an important role in formulating first impressions towards a website. Since a combination of different factors leads to a well-designed website, designers should consider how users gaze at the website. For example, as increased fixation duration reduces the possibility to perceive a website as high appealing, designers should avoid content that is likely to make the users' focus more on specific parts, such as a lot of text next to a picture of a product.

6.2. Limitations and Future Work

As with all studies, there are some limitations. First, we adopt a subjective approach, based on

8. References

- [1] M. Seckler, K. Opwis and A. N. Tuch, "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 (2015), pp. 375-389.
- [2] M. Moshagen and M. T. Thielsch, "Facets of visual aesthetics", *International Journal of Human-Computer Studies*, 68 (2010), pp. 689-709.
- [3] A. N. Tuch, J. A. Bargas-Avila and K. Opwis, "Symmetry and aesthetics in website design: It's a man's business", *Computers in Human Behavior*, 26 (2010), pp. 1831-1837.
- [4] T. Lavie and N. Tractinsky, "Assessing dimensions of perceived visual aesthetics of web sites", *International journal of human-computer studies*, 60 (2004), pp. 269-298.
- [5] E. Michailidou, S. Harper and S. Bechhofer, *Visual complexity and aesthetic perception of web pages, Proceedings of the 26th annual ACM international conference on Design of communication*, ACM, 2008, pp. 215-224.

questionnaires, in measuring perceptions of visual aesthetics [2]. Future studies, should combine the subjective and objective approach, which is based on screen design factors [1], along with eye-tracking techniques that will offer a better understanding of user behavior related with aesthetics. Second, the sample of the study is relatively small, however the low-level data we capture allow us to gain useful insight. Nonetheless, we plan to repeat this study, to increase the sample, and to verify and improve our findings. Also, in our study we included only experienced users in online shopping, however future studies should other types of customers, and differentiate based on experience, or based on gender [10]. Finally, to understand better the users, complexity theory should be combined with fuzzy-set qualitative comparative analysis (fsQCA) as it offers deeper insight into the data and can identify asymmetric relations within a sample [29, 30].

7. Acknowledgments



This project has received funding from the European Union's Horizon 2020 research and innovation programme, under the Marie Skłodowska-Curie grant agreements No 704110 and No. 751550.

- [6] M. Reimann, J. Zaichkowsky, C. Neuhaus, T. Bender and B. Weber, "Aesthetic package design: A behavioral, neural, and psychological investigation", *Journal of Consumer Psychology*, 20 (2010), pp. 431-441.
- [7] A. Dimoka, R. D. Banker, I. Benbasat, F. D. Davis, A. R. Dennis, D. Gefen, A. Gupta, A. Ischebeck, P. Kenning and P. A. Pavlou, "On the use of neurophysiological tools in IS research: Developing a research agenda for NeuroIS", *MIS quarterly*, 36 (2010), pp. 679-702.
- [8] S. Djasasbi, M. Siegel and T. Tullis, "Generation Y, web design, and eye tracking", *International journal of human-computer studies*, 68 (2010), pp. 307-323.
- [9] S. Djasasbi, M. Siegel, T. Tullis and R. Dai, *Efficiency, trust, and visual appeal: Usability testing through eye tracking, System Sciences (HICSS), 2010 43rd Hawaii International Conference on*, IEEE, 2010, pp. 1-10.
- [10] Y. M. Hwang and K. C. Lee, "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* (2017), pp. 1-10.

- [11] Q. Wang, S. Yang, M. Liu, Z. Cao and Q. Ma, "An eye-tracking study of website complexity from cognitive load perspective", *Decision support systems*, 62 (2014), pp. 1-10.
- [12] R. Arnheim, "Visual dynamics", *American Scientist*, 76 (1988), pp. 585-591.
- [13] M. Livio, *The golden ratio: The story of phi, the world's most astonishing number*, Broadway Books, 2008.
- [14] M. A. Just and P. A. Carpenter, "A theory of reading: From eye fixations to comprehension", *Psychological review*, 87 (1980), pp. 329.
- [15] O. Palinko, A. L. Kun, A. Shyrovkov and P. Heeman, *Estimating cognitive load using remote eye tracking in a driving simulator. Proceedings of the 2010 symposium on eye-tracking research & applications*, ACM, 2010, pp. 141-144.
- [16] L. P. Prieto, K. Sharma, L. Kidzinski and P. Dillenbourg, "Orchestration Load Indicators and Patterns: In-the-wild Studies Using Mobile Eye-tracking", *IEEE Transactions on Learning Technologies* (2017).
- [17] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka and J. Van de Weijer, *Eye tracking: A comprehensive guide to methods and measures*, OUP Oxford, 2011.
- [18] A. Smit and J. Gisbergen, "A short-latency transition in saccade dynamics during square-wave tracking and its significance for the differentiation of visually-guided and predictive saccades", *Experimental Brain Research*, 76 (1989), pp. 64-74.
- [19] W. S. Murray and A. Kennedy, "Spatial coding in the processing of anaphor by good and poor readers: Evidence from eye movement analyses", *The Quarterly Journal of Experimental Psychology*, 40 (1988), pp. 693-718.
- [20] R. Möller, *Perception through anticipation. A behaviour-based approach to visual perception, Understanding representation in the cognitive sciences*, Springer, 1999, pp. 169-176.
- [21] G. Lindgaard, G. Fernandes, C. Dudek and J. Brown, "Attention web designers: You have 50 milliseconds to make a good first impression!", *Behaviour & information technology*, 25 (2006), pp. 115-126.
- [22] A. Poole and L. J. Ball, "Eye tracking in HCI and usability research", *Encyclopedia of human computer interaction*, 1 (2006), pp. 211-219.
- [23] J. M. Henderson, *Visual attention and eye movement control during reading and picture viewing, Eye movements and visual cognition*, Springer, 1992, pp. 260-283.
- [24] E. D. Reichle, K. Rayner and A. Pollatsek, "The EZ Reader model of eye-movement control in reading: Comparisons to other models", *Behavioral and brain sciences*, 26 (2003), pp. 445.
- [25] R. Engbert, A. Nuthmann, E. M. Richter and R. Kliegl, "SWIFT: a dynamical model of saccade generation during reading", *Psychological review*, 112 (2005), pp. 777.
- [26] P. J. Unema, S. Pannasch, M. Joos and B. M. Velichkovsky, "Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration", *Visual cognition*, 12 (2005), pp. 473-494.
- [27] G. T. Buswell, *How people look at pictures*, University of Chicago Press Chicago, 1935.
- [28] A. Bronstein and C. Kennard, "Predictive eye saccades are different from visually triggered saccades", *Vision Research*, 27 (1987), pp. 517-520.
- [29] I. O. Pappas, P. Mikalef and M. Giannakos, *User Experience in Personalized E-Commerce: A Configurational Approach, 22nd Americas Conference on Information Systems (AMCIS)*, AIS, San Diego, USA, 2016.
- [30] I. O. Pappas, P. Mikalef, M. N. Giannakos and P. A. Pavlou, *Value co-creation and trust in social commerce: An fsQCA approach, 25th European Conference on Information Systems (ECIS)*, Guimarães, Portugal, 2017.

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).