

Using Biometrics to Evaluate Visual Design

Jane Lisa Dekker

Art Department, Northern California Valley State University

UXAD 272: Strategic Web Design

Professor Juan Liu, PhD

January 29, 2020

Using Biometrics to Evaluate Visual Design

A vast amount of research has been conducted regarding the importance of visual design, and its role as a mediator of a user's experience when browsing a site or interacting with an interface. In the literature, visual design is one aspect of website quality. Jones and Kim (2010) define website quality as "the perceived quality of a retail website that involves a [user's] perceptions of the retailer's website and comprises consumer reactions towards such attributes as information, entertainment/enjoyment, usability, transaction capabilities, and design aesthetics" (p. 632). They further examined the impact web quality and retail brand trust has on purchase intentions. Additional research examining e-commerce sites has shown web quality has an impact on both initial and continued purchase intention (Kuan, Bock, & Vathanophas, 2008), as well as consumer satisfaction (Lin, 2007). Moreso, research on the relationship between visual design and perceived usability (Stojmenovic, Pilgrim, & Lindgaard, 2014) has revealed a positive correlation between the two. As users' ratings of visual quality increase, their ratings of perceived usability follows a similar trend. Although this research spans various domains, the reliance on self-report measures to gauge concepts like visual design and web quality is prevalent throughout much of the literature.

Although some self-report scales are validated within the literature, there are still issues with the use of self-report questionnaires. One is the reliance on the honesty of the participant. This tends to be more of an issue in studies related to questionnaires that measure characteristics of the participant, rather than objective stimuli. More relevant to this study is the issue of introspection and memory. Surveys are often distributed after a task is completed, and its accuracy is dependent on the ability of the participant to remember their experience during the study. Multiple research studies have shown that human memory is far from static. This can

be dangerous if a researcher chooses to solely rely on self-report methods to test a hypothesis. We believe these self-report methods in tandem with biometric methods can help ensure the validity of the questionnaires, and provide information beyond the scope of self-report scales.

Research Questions

We know from previous research that the quality of websites mediates many aspects of e-commerce, and provides insight as to how consumers view the webpages in general. However, simply knowing a webpage is perceived as lower quality doesn't give insight as to what aspects of a page are disliked by a user. Additionally, it's possible that the user is misremembering aspects of the webpage or being dishonest in their assessment. Using eye tracking metrics, galvanic skin response, and facial expression measures in tandem with a scale aimed at measuring visual design quality has a couple of identifiable benefits. Using both can potentially identify patterns amongst the biometric measures and the questionnaire, which would strengthen the validity of the results. More so, the eye tracking data has the potential to identify patterns amongst websites of lower or higher quality.

If found, these patterns can be used to evaluate particular aspects of a page that are impacting the quality of a webpage. Overall, we are interested in answering two questions:

Research Question 1: Can attitudinal changes regarding substantial website redesigns be captured using biometric measures?

Research Question 2: How do biometric measures correlate with self-reported measures of visual appeal?

Answering these questions has the potential to provide a method of justification for design changes, ranging from minor tweak to complete rebrands. There is not an easy way for companies to quantitatively analyze visual design decisions. A method for doing so would help companies evaluate visual designs before implementation in order to cost-justify them. To this end, we hope to demonstrate that biometric measurements can be used with questionnaires to verify and validate potential design changes a company or organization might want to implement.

Conclusion

By examining data from test subjects during a brief exposure to several websites, we hoped to explore the relationship between the self-reported evaluation of visual design quality and key biometric measurements of a subject's emotional valence and arousal. Subjects were exposed to ten pairs of websites before and after a substantial visual design change and asked to evaluate the website based on their initial impressions of the site's visual design quality using the VisAWI-S scale, as shown in Table 1.

During this assessment we collected GSR, facial expressions (limited by errors in initial study configuration), pupillary response, and fixation data using iMotions software coupled with a Tobii eye tracker, Shimmer GSR device, and Affdex facial expression analysis toolkit. This data was analyzed, in Table 2, to discover relationships between the independent and dependent variables, as well as relationships between certain dependent variables.

References

- Jones, C., & Kim, S. (2010). Influences of retail brand trust, off-line patronage, clothing involvement and website quality on online apparel shopping intention: Online apparel shopping intention. *International Journal of Consumer Studies*, 34(6), 627–637.
<https://doi.org/10.1111/j.1470-6431.2010.00871.x>
- Kuan, H.-H., Bock, G.-W., & Vathanophas, V. (2008). Comparing the effects of website quality on customer initial purchase and continued purchase at e-commerce websites. *Behaviour & Information Technology*, 27(1), 3–16. <https://doi.org/10.1080/01449290600801959>
- Lin, H.-F. (2007). The impact of website quality dimensions on customer satisfaction in the B2C e-commerce context. *Total Quality Management & Business Excellence*, 18(4), 363–378.
<https://doi.org/10.1080/14783360701231302>
- Stojmenovic, M., Pilgrim, C., & Lindgaard, G. (2014). Perceived and objective usability and visual appeal in a website domain with a less developed mental model. *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: The Future of Design*, 316–323. <https://doi.org/10.1145/2686612.2686660>

Appendix

Table 1

Items included in the Vis-AWI-S instrument

Factor	Item
_____	_____
Simplicity	Everything goes together on the site.
Diversity	The layout is pleasantly varied.
Colorfulness	The color composition is attractive
Craftsmanship	The layout appears professionally designed
Familiarity*	I am familiar with this website

Note. Participants were asked about agreement with the item using a 7-point likert scale.

* question is simply to gauge familiarity for the study, and is not part of the Vis-AWI-S instrument

Table 2*Descriptive Statistics, Mean Difference, and p-values for Website Stimuli*

Website	Before		After		Mean Difference	p
	M	SD	M	SD		
Joy Kitchen	3.49	1.30	5.61	0.93	2.12	0.00
Seacom	3.27	1.59	5.35	1.20	2.08	0.00
Food Blog	3.59	1.30	5.59	0.80	2.00	0.00
Credit Union	3.29	1.26	5.18	1.07	1.89	0.00
Travelers	3.61	1.39	5.38	1.24	1.78	0.00
Sporcle	4.23	1.23	2.45	1.12	-1.78	0.00
Eagle	3.93	1.47	5.45	0.82	1.52	0.00
Oberlin	4.00	1.25	5.47	0.84	1.47	0.00
Valve	3.88	1.56	5.10	1.42	1.22	0.00
Hospital	4.47	1.33	5.48	0.85	1.01	0.00
Travel Blog	4.71	1.23	5.69	1.01	0.98	0.00
Space	4.35	1.55	5.29	1.09	0.94	0.00
School	5.04	1.44	5.63	0.80	0.60	0.06
Book Publisher	5.12	1.27	5.63	1.17	0.51	0.10
Sneakers	4.78	1.37	5.20	1.34	0.42	0.14
Stance	5.08	0.88	5.41	0.95	0.33	0.09
City	4.79	1.18	5.12	0.88	0.32	0.07
IEEE	3.95	1.30	4.26	1.40	0.31	0.24
Rise	5.08	1.00	4.89	1.27	-0.18	0.30
Audio Technica	3.94	1.52	4.05	1.37	0.11	0.71

Bloomberg	3.63	1.35	3.52	3.52	-0.11	0.73
-----------	------	------	------	------	-------	------

Note. Stimuli are ranked by largest to smallest absolute mean difference.