
Love at First Sight: Using Eye-tracking to Design for Visceral Appeal

Satyendra Nainwal

Office Experience Group
Microsoft Corporation
Redmond, WA, U.S.A
satynain@microsoft.com

Abstract

First impressions of visual appeal are an important factor that influence and predict overall desirability,

perceived usability and consequent success of a product. Despite their importance, there is a surprising lack of research methods and real-world case studies on how practitioners can better understand the relationship **between specific design elements of an interface and people's early perceptions of the interface.**

We propose a method that combines eye tracking analysis at the aggregate and individual level, first impressions testing, and qualitative feedback to compare people's first impressions of different designs. Our method provides **visual and actionable** feedback

to a design team allowing them to design for specific affective goals.

Author Keywords

Eye-tracking; visual design; desirability; visceral reactions; aesthetics; first impression; affect; emotions;

ACM Classification Keywords

H H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces, Evaluation/Methodology, Screen Design, User-centered Design.

General Terms

Human Factors; Design; Measurement.

Introduction

Designing a product or application that is desirable and pleasing to its target users is something every product team aspires to. There's been an increased focus on understanding visual aesthetics and their effect on user perceptions and the success of a product. Norman postulated that one tractable way to design towards specific goals was to target the three different levels at which people cognitively and emotionally form perceptions of a product: the visceral, behavioral and the reflective[17].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI'13, April 27 – May 2, 2013, Paris, France.

Copyright 2012 ACM 978-1-4503-1952-2/XX/XX...\$10.00.

To sum up, first impressions of a design are crucial to understand and get right, because:

1. Users can form rapid opinions and perceptions of a product/design based on first impressions.
2. These opinions are reliable and stable across time and influence other judgments like perceived usability and perceived information quality.
3. These influence subsequent judgments of the product due to:
 - a) Attention decrement interpretation.
 - b) Discounting interpretation.
 - c) Biased assimilation interpretation.

In Norman's model, the visceral response is the most immediate level where the reaction is in response to sensory aspects of a product that we perceive before we meaningfully interact with it. This is also the start of affective processing. The behavioral is the next level of processing that deals with perceptions we form as a result of our expectations and interactions with a product. And reflective is the third level which is intellectually induced and shaped by conscious reflection on past experiences. Designing for visceral reactions thus involves designing for affect through the senses. While this may include visuals, motion, sound, smell and touch for this paper we will limit the discussion to visuals of the interface.

It is important here to clarify a point about the terminology involving visceral. In Norman's view visceral reactions of visual appeal are physiological, genetically programmed responses devoid of reasoning[17,18]. Hassenzahl debates this notion, instead proposing that beauty judgments are "cognitive elaborations of the initial diffuse reaction" to stimuli"[5]. What both do agree upon is that the first response to visual stimuli induces a rapid, physiological reaction which is followed by a more expectation induced cognitive appraisal that leads to judgments of the visual.

This paper focuses on how we can understand the relationship between people's early reactions and judgment of the visual design and design elements they see. We first describe the theoretical basis and related work that inspired this approach. Next, we go on to show the need for this method especially in a fast paced iterative design cycle for a product like Microsoft Office. This requires quick meaningful feedback that

designers can feel engaged with and use to design iteratively. Then, we proceed to explain the methodology and setup we used. Finally we demonstrate how we've successfully used this method in the visual design of Office 2013 and lessons learnt that may help other practitioners.

RELATED WORK

Aesthetics, Perceptions and Preference

As our understanding of aesthetics grows, so does our realization of its importance in our everyday lives whether conscious or otherwise. Beauty has long been considered a basic requirement of good architecture[8], physical appearances are known to affect our social interactions[1], aesthetics was found to play an important role in new product development, marketing strategies, and the retail environment[20][10] and is a potent source of strategic competitive advantage[7] as well as an unquestionable determine a product's market success[2].

In the realm of interface design, Kurosu and Kashimura found a high relationship between users' judgments of the interface's aesthetics and 'apparent' (before use) usability[9]. Further research corroborated this finding removing both methodological as well as cultural bias and confirmed a strong correlation between users' aesthetic perceptions of the interface and the usability of the entire application[22]. Recent research also suggests that visual aesthetics of an interface is a strong determinant of users' satisfaction and pleasure[10] and, in fact, can improve their acceptability, learnability, comprehensibility and productivity [15]. All this affects the overall value people perceive in an application. Experiments have

Motivation: Understanding both what people see as well as **how what they see influences their perception** is crucial to using that understanding and applying it towards design iterations.

1. It allows for a deeper understanding of the relationship between design elements on screen and people's visceral reactions to the overall design.
2. Understanding this relationship provides the design team with **actionable** feedback for redesigns. The more direct and obvious this relationship is, the more engaging the team will be in acting on the feedback.
3. This targeted redesign allows for structured, iterative improvements of designs towards more positive visceral reactions.

shown that a website previously found visually very appealing was equally highly valued despite having a low usability[11].

Importance of First Impressions

People can form perceptions in response to visual stimuli extremely quickly. Research that shows that preferences to stimulus can be developed in 1-5 ms[25]. Some neurophysiological research supports this suggesting that emotional reactions can occur pre-attentively through a bundle of neurons that lead directly from the thalamus to the amygdala allowing for direct pre-cognitive responses to sensory stimuli[12].

In terms of interfaces, researchers have shown that people can form first impressions of a web page interface in as early as 50 milliseconds[12]. First impressions are an important influencer and predictor of overall appeal and perceived usability and are consistent and stable over time[23]. First impressions of visual appeal may also affect other qualities like information quality[4].

Initial impressions not only impact early reactions but have repeatedly been shown to strongly influence later interpretation. Numerous studies have shown this primacy effect where "information presented early in a sequence has more influence on final judgments than information presented late". [21] This is true in a variety of contexts ranging from Personality of people, art and paintings, to the perceived usability and aesthetic appeal of products, web sites and even graphs. [1][12][19] Molnar further goes on to state that first impressions are more indicative of pure aesthetic judgment[14] which makes them even more

important to understand the aesthetic appeal of an interface.

Tracking and understanding what people see

There have been attempts to study visual processing by tracking eye movement as early as 1930 if not earlier. One of the most famous studies to track eye movement and understand how it relates to perception was conducted by the Russian scientist Yarbis in 1967. Tracking eye movement of how people viewed Repin's painting "They did not expect him", one of Yarbis' conclusions was that the task people are performing has an implication on resulting eye fixations. Molnar further refined this framework where he differentiated between targeted scanning aimed at seeking knowledge vs. exploration aimed at just seeking pleasure[14]. Locher et al. provide further evidence where people first perform a 2 second global sweep of the image and then look at portions the image in finer detail. [13][16]

More recent work has focused on studying how the eyes move and perceive differences in styles, aesthetics and complexity of visual stimuli and how computational models can explain them[24].

Motivation for this work

Research corroborates what designers and artists have long known intuitively, that specific design features, such as movement, pattern and proximity, can influence visual processing[19] and hence people's first impressions. However, for designers to actually use this information they need to understand **both** what people see as well as how what they see influences their perception. Understanding this **relationship** is crucial for designing for desirable affective reactions by using that feedback in subsequent design iterations.

Summary of the Procedure

1. Participants see images of interfaces for an interval of three seconds.
2. They just observe this image with the goal to explore. (vs. a task goal)
3. The first couple of images would be example images to give them a sense of the time and the kind of feedback that would be useful.
4. The images were randomized by the eye-tracker.
5. They were told not to think out loud while they were watching the image. Once the image appeared and disappeared, there would be a black screen during which they could talk about their reactions to the image.
6. Before the next image was presented the researcher would ensure participants are ready and their gaze is fixed at the center of the screen.

I set out to take the learning from the existing research and build on it to help our team better understand the relationship between visual elements and people's perceptions so that we can design the visuals to evoke specific affective outcomes.

Methodology and Setup

We had a total of 9 participants in the study for a one on one study. The composition was 5 Information Workers and 4 students with an age range of 21 years – 55 years and a mean of age of 30.4 years. The participants were recruited based on their answers to a screener on Office usage. All 9 of them were familiar with Office 2010 and either used it as their primary productivity suite or in conjunction with another productivity application. The participants were informed that for this set of exercises they would not need to use the mouse or keyboard. They were then given the instructions summarized in the left pane.

Later, they had an opportunity to look at all the designs more closely for a much longer time, provide detailed feedback and fill a custom semantic scale questionnaire. We will not report on the detailed feedback they gave in the later part of the study in this paper, to focus the discussion on first impressions. Finally, we set up the experiment so that a couple of designs came up more than once for both the older and the new designs. This was to put a check in place to verify that people were being consistent with their feedback.

Setup

We used a Tobii T60 table-mounted eye tracker which is a binocular eye tracker integrated in a 17 inch TFT flat panel monitor. We set the resolution to 1280*1024 for all participants with a sampling rate of 60 Hz.

Results

Unlike most other experiments where the visual stimuli were images, paintings, or web pages with more engaging and meaningful content we had an interface dominated by buttons and tabs. Interestingly users immediately found differences within the interfaces. In the couple of the designs that we had repeated we got consistent responses and often got the feedback that people felt it was very similar to a previous design they had seen.

In the immediate feedback following the 3 second glimpse, the new designs did better than Office 2010 designs for all applications. For some applications the preference was stronger than the others. For the new Outlook designs in particular, all participants strongly preferred the new designs compared to Outlook 2010.

At the end of the eye tracking part of the study we showed participants images of both versions of the Office Suite - Word, Excel, PowerPoint, Outlook and OneNote. For the graphs below Suite A refers to Office 2010 and suite B to the new designs.

The participants were free to look at both suites, compare and go back and forth between them as many times as they pleased (no time limit). Participants were asked them to rate each suite on a semantic scale along a few attributes. Their perceptions after the detailed look correlated very closely to the feedback they provided during the 3 second glimpse. Keywords like Clean and Personal that participants had spontaneously used most frequently in the 3 second glimpse were rated highly in the detailed look too. Here are the results from 4 of those attributes. Interestingly aesthetics did influence their perception of

Participant reactions after 3 second look:

Outlook 2010: *"It was really busy and I didn't really know where I should be looking. My eyes darted directly to the photo of the person that was on the right. Ummmm that caught my eye, but the rest of the page I wasn't really sure where to look. It was very cluttered. And I felt.....overwhelmed a little bit."*

New Outlook Design: *"I really liked this one because the bottom where it had the messages and calendar I like how the lettering was very big but it wasn't loud because of its softness and the soft orange color and picture I liked. I really liked how it looked organized and just had a few columns and again with the curves and the softness of the lettering it was pleasing to the eye and pleasing to look at. Especially because the separations between each column were subtle but they were still there which I liked".*

"ease of use". The greatest difference between the suites was on the attributes of clean and personal. The qualitative feedback indicated that the clean was driven by a feeling of cleaner and fewer lines, flatter surfaces and more whitespace. The personal feedback was mostly around the color being more prominent and attractive as well as pictures of people, and the information feeling less overwhelming.

Eye Tracking Analysis

In addition to people's reported reactions to the designs we were also interested in understanding what design elements were causing those reactions. So we first analyzed the aggregate heat maps and then each individual's eye tracking behavior. Outlook was one application where participants' reactions were most polarized for the two designs, so we'll present that as example to illustrate the method.

In the gaze opacity maps below the darker the area, the lesser the aggregate gaze fixation and the whiter the area the more the gaze fixation on that area. Visualizing it in this way had a great impact with the team. In addition to seeing what people focused at in the first 3 seconds, the team could also get a sense what areas of the image the foveal vision didn't focus on in the first 3 seconds!

The most prominent fixation areas for the new designs were the picture of the sender and the orange colored "File" tab, along with the area towards the center of the application, which was the area where the participants' gaze was initially focused to start with.

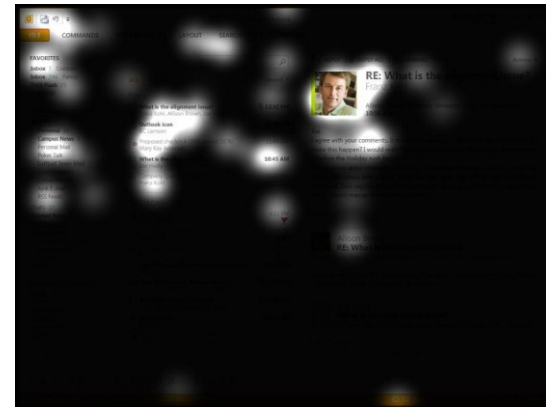


Figure 1: New Outlook design: **perceived as clean and personal**; N=9 users, time= 3 s

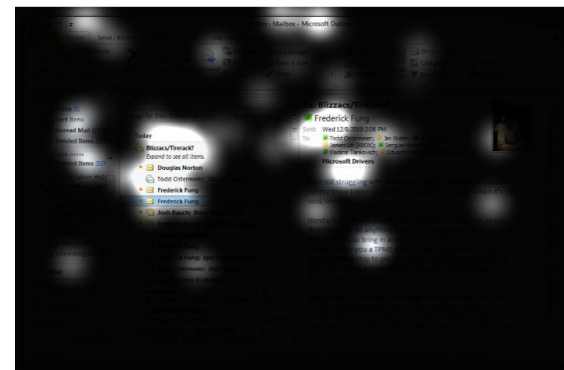


Figure 2: Outlook 2010: **Perceived as cluttered**; N=9 users; time=3 s

For the Office 2010 designs, the most prominent gaze fixation areas are around the center, where the gaze was initially focused and then the "reading pane" area to the right of the initial fixation. The gaze fixations

Figure 3: Aggregate heat map evolution for Outlook 2010 for all 9 participants. From left to right the three images show the heatmap at: **t < 500 milliseconds, t=1 second, t= 3 seconds.**

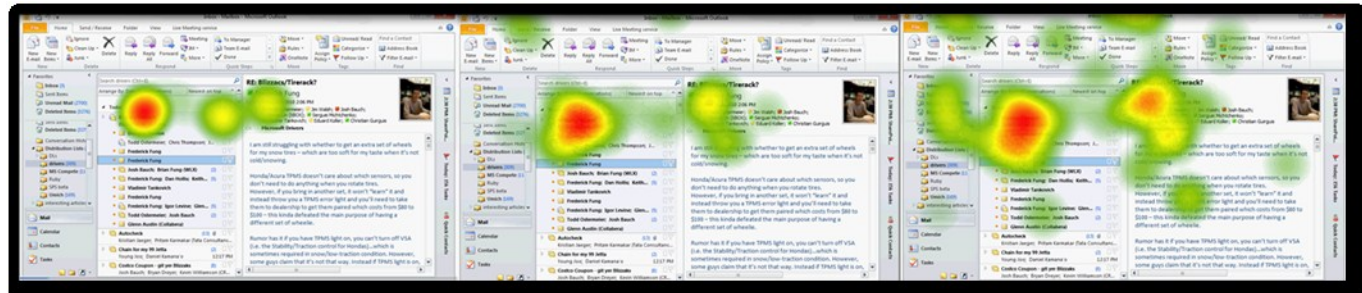
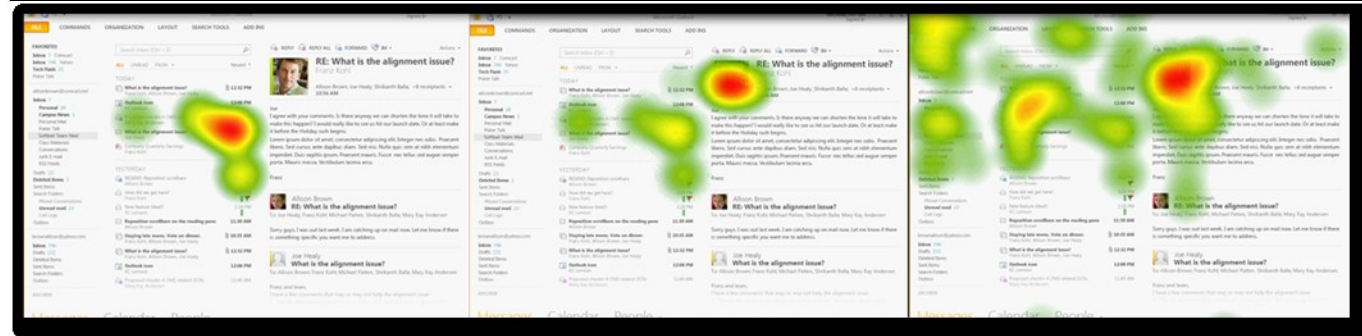


Figure 4: Aggregate heat map evolution for new Outlook designs for all 9 participants. From left to right the three images show the heatmaps at: **t < 500 milliseconds, t=1 second, t= 3 seconds**



here appear more scattered and contain mostly text heavy regions.

To understand the order of the gaze fixations we looked at a gaze plots and then created a video of the aggregate heatmap that would show the movement of gaze fixation at the aggregate level for those 3 seconds. Figures 3 and 4 below show snapshots from that video at t < 500 milliseconds, t=1 second and t=3 seconds for all 9 users. From the gaze plot numbers and the aggregate heat map now the story starts to get clearer!

For Outlook 2010 we can see that at the aggregate level the gaze starts at around the upper-center of the application stays there for a while (red hotspot), then

moves to the right, scans around the area, and either moves back or scatters. For the new Outlook design, gaze again starts at the center, then uniformly moves to the picture of the sender (red spot now shifts to picture), scans around the area and then starts moving to the top left to the File Tab. Also notice the smoother transitions between the hot-spots for the new design compared to Office 2010.

What do people see at the individual level?

From the aggregate patterns we looked at in the last section we can conclude that, at the aggregate level, the new Outlook designs have clear, consistent hot spots where the fixations are focused. We then looked at a finer level of granularity to understand what each person saw at the individual level. At the individual

level too, We found that for the new Outlook designs each of the 7 participants' eyes were guided not just to the same elements but also **guided to them in a similar order!** For the new designs each of the participants' eyes starts from the center of the screen. The eye is then guided first to the picture of the person who sent the email as the first prominent fixation and then the "File" button as the next prominent fixation. This can be clearly seen in the individual gaze opacity maps in Figure 5 of four example participants. Each of their eyes is guided to the same design elements in the same order. **Thus the visual design creates a clear information hierarchy by guiding the eyes consistently to specific design elements.**

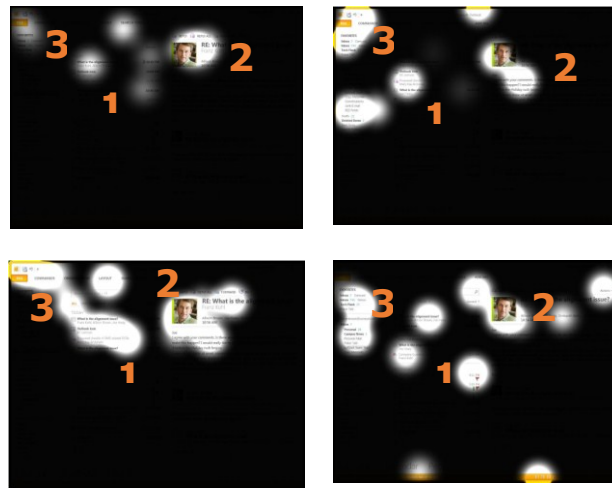


Figure 5: Individual Gaze opacity maps; New Outlook Designs

Compare that with the individual gaze patterns for Outlook 2010 designs in Figure 6. In the Outlook 2010 designs there are no consistent areas of fixation

and there is no coherence in the order in which the eyes move.

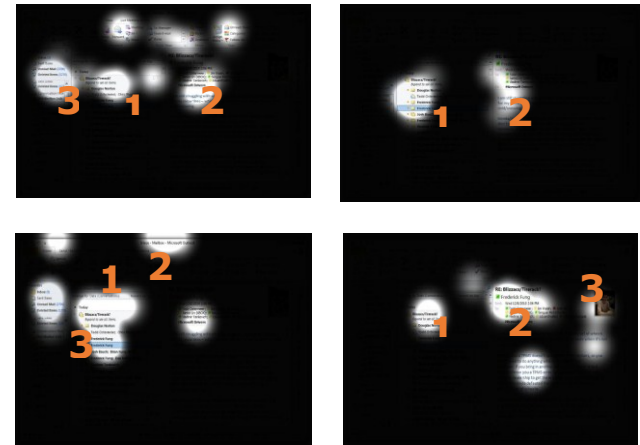


Figure 6: Individual Gaze opacity maps; Outlook 2010 Designs

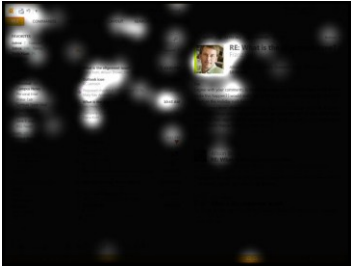
The visual design here does not create a good information hierarchy and hence does not guide the eye. Therefore, individual participants' eyes focus on different areas and do so in no predictable manner.

Discussion of results

Triangulating the qualitative feedback of people's reactions, both to the 3 second glimpse as well as after the study, along with a detailed analysis of the eye tracking data both at the aggregate and individual level we confirmed some of the results from existing literature and also found interesting new insights that we can build upon in the future.

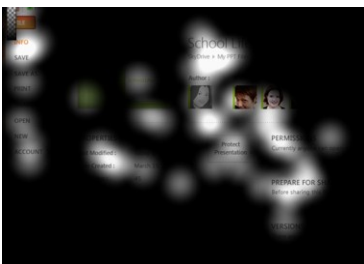
We confirmed that people can form opinions about the overall appeal of an interface even in a quick glimpse –

Generating Insights for Reuse



The sender's picture contributed to strong perceptions of "personal" in the Outlook study (above).

We used this insight as a hypothesis to design a different application interface with prominent pictures of people (below) which again scored high on personal compared to a design without that element



3 seconds in this case. Further, these perceptions correlated closely with their opinion after exploring the interface at length. Our study also confirmed that perceptions of overall appeal can affect a-priori perceptions of ease of use. However, the triangulation of eye tracking data, aggregate and individual, also taught us things we didn't find in existing literature. We'll discuss those in a little more detail below.

Importance of triangulating what people see with what they say

We know the first response to visual stimuli induces a rapid, physiological reaction which is followed by a more expectation induced cognitive appraisal that leads to judgments of the visual. While self-reported qualitative feedback has been proven successful and useful for understanding people's overall impressions of the designs, it is not a reliable way of understanding what led to those impressions. In fact **in some cases, it may actually be misleading**. For example in analyzing the Outlook example above, the participant quoted told us that her eyes "*darted directly to the photo of the person that was on the right*". Without eye tracking data we would conclude that she first looked at the photo. However, triangulating with the eye-tracking data we found that the person's photo was in fact the **last element she saw**. The fact that she recalled this as something she saw first is likely because of the serial position effect. She was in free recall mode and hence relying on short term memory and recalled the latest experience first. So **triangulating what people say with what people see is important if we want reliable data on why certain designs might lead to certain perceptions using retrospective think aloud**.

Relation between design elements and affective responses

The outputs from our method can show relationships between eye gaze patterns and interface elements. This allows us to generate, compare patterns and get a deeper understanding of the relationship between what people see and what they say. Norman states that the visceral level operates by pattern matching[17]. That implies that the detection of a certain pattern will trigger a predictable class of affective responses. Some examples he cites that trigger positive reactions include smiling faces, "attractive" people, rounded smooth objects. We heard similar feedback for the new Outlook designs. For example several people reported the photo of the sender being a reason for rating the new design high on the "personal" attribute. Eye tracking confirmed it was a prominent gaze fixation area in the new interface and one of the first elements people saw.

Generating insights, hypothesis and Reuse

Understanding the relationship between design elements and affective responses can lead to generation of new insights and hypothesis. For example, learning that having pictures of people along with flatter surfaces and good balance of whitespace can lead to perceptions of clean and personal in one interface can lead to insights for designers on how they can reuse this in other places in a similar context. Of course, this is just a hypothesis that will need to be tested and iterated on but it does provide a powerful starting point.

Visual, actionable feedback

Another practical learning from working very closely with the visual design team has been that the medium and manner in which the results are communicated to

the designers has a large impact on how engaged they are in the process what happens with those results.

We learnt that indirect ratings like graphs, rating-scores and scorecards alone to show how appealing a design wasn't very engaging. However, having a visual representation of what peoples saw, and how that might have affected their perception had the entire team very engaged and eager to participate. The heatmaps and gaze plots led to brainstorming sessions after which the team would go back create variations of the designs excited to see if they could guide the way people saw the designs.

Future work

We're continuing to explore and refine our methods. Two issues that we're specifically looking at, not explored here are different methods to collect the qualitative feedback with the eye tracking and how that affects results. We'd also like to correlate the findings here with behavioral and visceral level feedback to get a more holistic view of designing products that delight not only at the initial level but also in longitudinal usage.

ACKNOWLEDGMENTS

I'd like to thank my managers Kent Lowry and Shawn Lipstein for encouraging me, despite tight deadlines, to explore methods where we weren't certain to get results. I'd also like to thank the other designers and program managers I worked with; in particular, John Schilling, Keri Talbot, Chris Edwards, Owen Braun and Krista Bendig for all their help.

REFERENCES

1. Anderson, N.H. and Hubert, S. Effects of concomitant verbal recall on order effects in personality impression formation. *Journal of Verbal Learning and Verbal Behavior* 2, 5-6 (1963), 379–391.
2. Bloch, P.H. Seeking the Ideal Form: Product Design and Consumer Response. *Journal of Marketing* 59, 3 (1995), 16–29.
3. Bornstein, R.F. and D'Agostino, P.R. Stimulus recognition and the mere exposure effect. *Journal of Personality and Social Psychology* 63, 4 (1992), 545–552.
4. Geest, T.V.D. and Dongelen, R.V. *What is beautiful is useful - Visual appeal and expected information quality*. IEEE, 2009.
5. Hassenzahl, M. The Interplay of Beauty, Goodness, and Usability in Interactive Products. *Human-Computer Interaction* 19, 4 (2004), 319–349.
6. Kim, H. and Fesenmaier, D.R. Persuasive Design of Destination Web Sites: An Analysis of First Impression. *Journal of Travel Research* 47, 1 (2008), 3–13.
7. Kotler, P. and Rath, G.A. Design: a Powerful But Neglected Strategic Tool. *Journal of Business Strategy* 5, 2 (1984), 16–21.
8. Kruft, H.W. *A History of Architectural Theory: From Vitruvius to the Present*. Zwemmer and Princeton Architectural Press, New York, 1994.
9. Kurosu, M. and Kashimura, K. Apparent usability vs. inherent usability. *Conference companion on Human factors in computing systems CHI* 95, (1995), 292–293.

10. Lavie, T. and Tractinsky, N. Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human-Computer Studies* 60, 3 (2004), 269–298.
11. Lindgaard, G., Dudek, C., Sen, D., Sumegi, L., and Noonan, P. An exploration of relations between visual appeal, trustworthiness and perceived usability of homepages. *ACM Transactions on Computer-Human Interaction* 18, 1 (2011), 1–30.
12. Lindgaard, G., Fernandes, G., Dudek, C., and Brown, J. Attention web designers: You have 50 milliseconds to make a good first impression! *Behaviour Information Technology* 25, 2 (2006), 115–126.
13. Locher, P., Krupinski, E.A., Mello-Thoms, C., and Nodine, C.F. Visual interest in pictorial art during an aesthetic experience. *Spatial Vision* 21, 1-2 (2007), 55–77.
14. Molnar, F. *About the role of visual exploration in aesthetics*. Plenum Press, New York, 1981.
15. Ngo, D.C.L., Teo, L.S., and Byrne, J.G. Modelling interface aesthetics. *Information Sciences* 152, 0 (2003), 25–46.
16. Nodine, C.F., Locher, P.J., and Krupinski, E.A. The Role of Formal Art Training on Perception and Aesthetic Judgment of Art Compositions. *Leonardo* 26, 3 (1993), 219.
17. Norman, D.A. Emotion and design: Attractive things work better. *Interactions*, July-Aug (2002).
18. Norman, D.A. *Emotional Design: Why We Love (or Hate) Everyday Things*. Basic Books, 2004.
19. Renshaw, J.A., Finlay, J.E., Tyfa, D., Ward, R.D., and Park, B. Designing for Visual Influence: an Eye Tracking Study of the Usability of Graphical Management Information. c (2003), 144–151.
20. Russell, J.A. and Pratt, G. A description of the affective quality attributed to environments. *Journal of Personality and Social Psychology* 38, 2 (1980), 311–322.
21. Tetlock, P.E. Accountability and the perseverance of first impressions. *Social Psychology Quarterly* 46, 4 (1983), 285–292.
22. Tractinsky, N., Katz, a. ., and Ikar, D. What is beautiful is usable. *Interacting with Computers* 13, 2 (2000), 127–145.
23. Tractinsky, N., Cokhavi, A., Kirschenbaum, M. and Sharfi, T. Evaluating the consistency of immediate aesthetic perceptions of web pages. *International Journal of Human - Computer Studies* 64, 11 (2006), 1071–1083.
24. Wallraven, C., Cunningham, D., Rigau, J., Feixas, M., and Sbert, M. Aesthetic appraisal of art - from eye movements to computers. (2009).
25. Zajonc, R.B. Feeling and thinking: Preferences need no inferences. *American Psychologist* 35, 2 (1980), 151–175.