AVOIDING AD AVOIDANCE: FACTORS AFFECTING THE PERCEPTION OF ONLINE BANNER ADS

Felix Portnoy

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of Information and Library Science.

Chapel Hill 2012

Approved by:

Dr. Todd Barlow

Dr. Robin Jeffries

Dr. Sri Kalyanaraman

Dr. Diane Kelly

Dr. Gary Marchionini

©2012 Felix Portnoy ALL RIGHTS RESERVED

ABSTRACT

FELIX PORTNOY: Avoiding Ad Avoidance: Factors Affecting The Perception Of Online Banner Ads

(Under the supervision of Dr. Gary Marchionini)

This dissertation examined the effect of search type, ad saliency, and ad repetition on the perception of online banner advertisements. In the first study, 48 student participants conducted simulated search tasks using mixed factorial design where search type (known-item vs. exploratory) was manipulated within-subject and the banner saliency level (low (black and white) vs. medium (color) vs. high (color animation)) was manipulated between subjects. The results showed a significant effect for search type, such that during an exploratory search task the participants had a higher average number of eye fixations on the banner ads compared with known-item search. In addition, there was a significant difference between high and low ad saliency levels, such that participants exposed to low salient ads had a higher average number of eve fixations on the banner ads as compared with high salient ads. There was no significant effect of ad repetition on ad perception. A second study replicated the original experimental design but used four novice Internet users. The results from the second study provide preliminary support to the asymptotic habituation model, which predicts an inverse decline of an orienting response to banner ads as a function of repetition. This dissertation concludes with applicable design recommendation for banner ad deployment to ensure visibility while maintaining a positive user experience.

TABLE OF CONTENTS

LIST OF	TABLES	vi
LIST OF	FIGURES	vii
ACKNOW	VLEDGEMENTS	xi
Chapter		
I.	PROBLEM DOMAIN	1
	Research Question	5
II.	LITERATURE REVIEW	6
	Ad Avoidance	6
	Theories of Attention and Mental Workload	25
	Physiological Measures in Human-Computer Interaction	45
	Summary	71
	Hypotheses	74
III.	STUDY 1—EXPERT USERS	75
	Experimental Design	75
	Results	87
	Summary	100
IV.	STUDY 2—NOVICE USERS	102

	Experimental Design	102
	Results	104
	Summary	107
V.	DISCUSSION	109
	Search Type Matters	109
	Less is More	111
	The Effect of Habitation on the Response to Banner Ads	113
	The Use of Physiological Measures	113
	Recommendations Based on the Research Findings	114
VI.	FUTURE WORK	118
VII.	CONCLUSIONS	121
APPEND	ICES	123
Appendi	x A – Search Tasks	123
Appendi	x B – Information Sheet	133
Appendi	x C – Consent Form	136
Appendi	x D – Recall Questionnaire	137
Appendi	x E – Recognize Questionnaire	138
Appendi	x F - Attitude Towards Ads Questionnaire	141
Appendi	x G - Web User Self Efficacy Scale (WUSE)	143
Appendi	x H – Information Sheet for Study 2	147
Appendi	x I – Consent Form for Study 2	150
REFERE	NCES	151

LIST OF TABLES

Table

1.	Summary of the investigation of the difference in orienting response to banner ads between known-item and exploratory search task	21
2.	List of common eye tracking metrics used in computer interface evaluation	52
3.	Sequences of search tasks (KI=known-item; E=exploratory)	.79
4.	Test of normal distribution	.90
5.	Model fit results for the effect of repetition on number of fixations	95
6.	Descriptive statistics for the Stroop test analysis (all units are in milliseconds)	99
7.	Model fit results for the effect of repetition on number of fixations based on novice users	104

LIST OF FIGURES

Figure

1.	An example of banner ad used by the New York Times	1
2.	An example of the banner ads used by Drèze & Hussherr (2003)	3
3.	An example of an aided recall question. On the left is the ad as it appeared on the web page. On the right is the same ad where the brand is masked. The participants were asked if they remembered seeing this ad and, if so, what was the name of the brand.	4
4.	Percentage of online advertising revenue generated by banner ads (IAB, 1997-2010)	4
5.	An example of stimuli used by Benway (1998)	6
6.	An example of the search task used by Burke et al. (2005). The sentence on the top was used as a precue that disappeared immediately as the list of headlines appeared	7
7.	Depiction of the asymptotic habituation model. Stimulus repetition leads to lower likelihood of response to that stimulus as a result of habituation	9
8.	A modified illustration of the Habituation-Tedium Theory	2
9.	Model of search activities (Marchionini, 2006)1	3
10.	A screen shot of the stimuli used by Pagendarm & Schaumburg (2001)1	6
11.	The website used by Calisir & Karaali (2008) and the three banner ad variations	8
12.	Geographical map used in the experiment by Van Orden et al. (1993)2	2
13.	An example of the stimuli used at Tullis et al. (2009)2	3

14.	Diagram of the information flow according to the filter theory (Moray, 1969, p. 29)	28
15.	Treisman's input selection model (copied from Moray, 1969, p. 31).	31
16.	Feature integration model (copied from Treisman & Gormican, 1988, p. 17)	33
17.	Search times as a function of type of task and display size (copied from Treisman & Gelade, 1980, p. 104)	34
18.	Basic components of the guided search 2.0 model (copied from Wolfe, 1994, p. 205)	35
19.	The structure of multiple resource theory (copied from Wickens, 2002, p. 163)	39
20.	Interaction between different levels of distractors susceptibility as a function of low and high cognitive and perceptual load	43
21.	Schematic diagram of the human eye. Source: Wikipedia. Retrieved from http://bit.ly/97RBPt	48
22.	Infrared deflection and bright pupil as seen by the infrared camera.	50
23.	The infrared light deflection glint allows the eye tracking system to locate the participants' focal point	50
24.	The stimulus used by Goldberg et al. (2002) showing the scanpath and the area of interest	55
25.	An example of a modern galvanic skin response sensor. Source: Qubit Systems. Retrieved from http://bit.ly/c4x2fK	58
26.	The static electrical "bath." In this version, an electrode over the head directed the discharge, producing static "wind" (Neumann & Blanton, 1970, p. 458)	59
27.	R Peaks of ECG Signal & RR Intervals. Source: http://bit.ly/c0UGA6	63
28.	An example for the stimuli used in Diao and Sundar	. 69

29.	Hardware setup in the interaction design lab	77
30.	Instruction/home page of the website	78
31.	"Search results" page for Bulgaria with known-item search tasks	80
32.	"Search results" page for Bulgaria with an exploratory search task	80
33.	An example of a New York Times article layout	81
34.	Example of sub-page design.	82
35.	The five banner ads used in this experiment.	83
36.	Illustration of the experimental procedure	84
37.	Example of a LookZone used throughout the data analysis	88
38.	An example of a manual gaze shift. The top image shows the original gaze data and the bottom image shows the adjusted gaze data. In this case, the gaze data was shifted 60 pixels to the right and 20 pixels upwards.	89
39.	Box plots of the number of fixations data	91
40.	Box plots of the <i>transformed</i> number of fixations data showing normal distribution	91
41.	The effect of ad saliency and search type on number of fixations. Error bars represent standard error.	94
42.	Distribution of the fixation data as a function of repetition (trial order).	95
43.	Frequency distribution (histogram) of the recall questionnaire scores.	98
44.	Distribution of the fixation data as a function of repetition among novice users.	105
45.	Depiction of the data analysis using Biopac Student PRO v.3.7.3	106
46.	Distribution of the galvanic skin response data as a function of repetition	107

47.	Distribution of the heart rate data as a function of repetition	107
48.	Heat map showing the fixation distribution in exploratory search task (left image) and known-item search task (right image)	110
49.	An illustration of the differing salience levels of the banner ads as a function of their surrounding objects	115
50.	An illustration of a low salient promotional message embedded in a financial webpage	117
51.	The Tobii X1 Light Eye Tracker allows the researcher to conduct eye-tracking studies in the field using a laptop	117

ACKNOWLEDGEMENTS

It has been 11 years since I walked into my first class in the University of Haifa. As a high school dropout, the thought that I would ever complete a bachelor's degree was laughable. Now, as I approach the completion of my doctoral degree, I remind myself of all those who have been there for me, and without whom I could not have accomplished this important milestone.

I would like to start by thanking my advisor, Dr. Gary Marchionini, who took a leap of faith and allowed me, under unusual circumstances, to join the School of Information and Library Science (SILS). Through his guidance, patience, and care, I was able to fulfill my scholarly potential. His resourcefulness and insights proved to be indispensible in the course of my doctoral studies.

Although I am listed as a sole author, this dissertation is the product of a collaborative effort by my doctoral committee members: Dr. Todd Barlow, Dr. Robin Jeffries, Dr. Sri Kalyanaraman, and Dr. Diane Kelly. I am grateful for the time they volunteered to provide insight and instructive critique of my research endeavors.

Special thanks to Google for providing the generous funding that allowed me to focus on my research interests and this dissertation.

To the SILS faculty, specifically Dr. Jaime Arguello, Dr. Bradley Hemminger, Dr. Robert Capra, and especially Dr. Barbara Wildemuth, who always welcomed me into their offices and provided a useful word of advice and encouragement.

To the SILS staff, specifically Scott Adams, Lara Bailey, Aaron Brubaker, Tammy Cox, Wakefield Harper, Mark Snyder, Susan Sylvester, Michelle Taylor, and Shaundria Williams: all have bent over backwards to help me accomplish my goals through the course of my studies.

To the SILS Ph.D. students, the most creative and talented bunch of people I have had the pleasure to encounter. Thank you for your help in the early stages of my experiments, whether by revamping my experimental procedure or by serving as guinea pigs in my lab. Your feedback was crucial for the success of this dissertation.

I would also like to thank my previous advisors, Dr. Micah Leshem of the University of Haifa and Dr. Paul Milgram of the University of Toronto, who indirectly contributed to this undertaking.

Dr. Leshem was the first professor who allowed me to work in a research lab during my freshman year. In the course of two years under his mentorship, I had a rare opportunity to learn how to conduct experimental studies. That experience sparked my curiosity and led me to pursue a research career.

Dr. Milgram not only taught me the fundamentals of human factors, but also inspired me by setting a personal example of attaining academic excellence. In addition to inviting me into his lab, he welcomed me into his home to celebrate holidays along with his family—a gesture that was much appreciated by a solitary foreign student.

On a personal note, I would like to thank my parents, Alex and Tatiana Portnoy, who can finally have some *nachas* after long years of endless care. To my parents-in-law, Richard and Janice Foleck, for treating me as their own son and for babysitting hundreds of hours to help me complete my work in a timely manner. To my sister and bothers-in-law Natalie Kalman, Yossi Kalman, and Dr. Adam Foleck for their encouragements.

To my cherished daughter, Amelia, for the delightful distractions and helping me realize my priorities in life. Finally, to my dearest wife, Erinn, for taking an active role in this journey as a reviewer, counselor, and most of all, friend. I love you •

PROBLEM DOMAIN

On October 27th, 1994, the novel online magazine HotWired (known today as Wired Magazine) introduced a new marketing tool that would revolutionize business and publishing—a banner ad for AT&T (D'Angelo, 2009; Zeff & Aronson, 1999). The banner ad is a type of graphic image that appears on a website. When users click on that image, they are redirected to another website that provides information about the product shown in the graphic image. In 2010, banner ads (similar to the highlighted ad shown in Figure 1) generated revenue of \$6.2 billion in the United States alone, representing a little over 25% of the online advertising revenue that year (IAB, 2011).

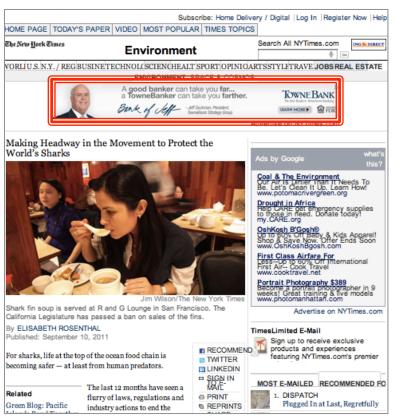


Figure 1. An example of banner ad used by the New York Times.

To assess the banner ad's effectiveness, advertisers commonly rely on a measure called click-through rate (CTR), which indicates the ratio between the number of page visits and the number of times the banner ad was clicked. While the CTR for the first banner ad was an astonishing 78%, the CTR has steadily declined to an average of <0.4% (Steinberg, 2011). This drastic decline led some publishers to conclude that the prevalent model of free content paid by ads was unsustainable, and led to a growing number of online publishers establishing paywalls to their sites (Clark, 2009; Peters, 2011).

However, several studies have suggested that CTR may not be an appropriate measure of ad exposure since ads affect users in the pre-attention levels, thus making ad memory tests more appropriate than CTR (Briggs & Hollis, 1997; Danaher & Mullarkey, 2003; Drèze & Hussherr, 2003; Shapiro, MacInnis, & Heckler, 1997). For example, Drèze & Hussherr (2003) used eye tracking to measure the number of fixation on mock banner ads embedded in a portal service (see Figure 2). Results obtained from the eye gaze analysis showed that the participants looked directly only at half of the banner ads during the study. When the participants were later asked whether they recalled seeing ads during their tasks, 46.9% of the participants responded "yes", though they were not able to correctly identify the ads that they were exposed to. In a follow-up study, Drèze & Hussherr (2003) examined how brand awareness changed as a function of mere exposure to ads.



Figure 2. An example of the banner ads used by Drèze & Hussherr (2003).

Prior to administering the study, they asked their participants to rate their familiarity with 10 product brands. The participants were then asked to browse through nine web pages while completing different tasks such as word entry in a search field. Some brands that the participants were asked about in the pre-test questionnaire appeared in those websites, while other did not. Using a post-stimuli questionnaire, the authors found that the awareness (familiarity ratings) of brands that were included in the study significantly increased by an average of 2.8% while the awareness of brands that were not included in the study did not change from the pre-stimuli ratings. In a recall question ("please enter the brand(s) advertised"), brands that appeared in the study had an average recall rate of 11.4% compared to only 2.47% for brands that did not appear in the study. In an aided recall question (see Figure 3), the average was even higher—30.1% of the respondents correctly recalled the brands that appeared in the study (no comparison was made with brands that were not included in the study). The authors concluded that the banner ads had a long-term effect on the users' memory and attitude towards the brands.

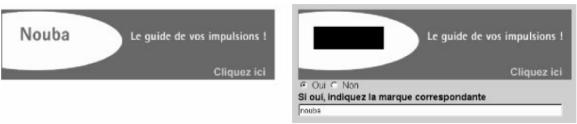


Figure 3. An example of an aided recall question. On the left is the ad as it appeared on the web page. On the right is the same ad where the brand is masked. The participants were asked if they remembered seeing this ad and, if so, what was the name of the brand.

Despite these and other results showing that study participants are affected by banner ads, the use of banner ads in the industry has declined sharply since the late 1990's when banner ads generated over 50% of online advertising revenue (see Figure 4).

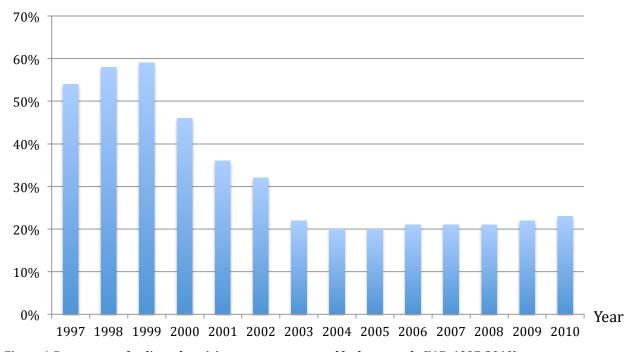


Figure 4. Percentage of online advertising revenue generated by banner ads (IAB, 1997-2010)

Research Question

In light of these issues, the goals of this dissertation are as follows:

- 1. Create a conceptual model that explains users' interactions with banner ads.
- Develop a set of design recommendations that will allow online publishers to create more effective banner ads with minimal disruption to their users' task flow.

To address these goals, my research endeavors are focused on the question: which factors affect users' perception of online banner ads?

At this time, based on the available research, it is unclear why the effectiveness of online banner ads was diminished. Hence, the purpose of this research is to examine what factors may affect the users' response towards banner ads and what factors could make banner ads more effective while maintaining a positive user experience.

LITERATURE REVIEW

Ad Avoidance

Ad avoidance is a recurring phenomenon in which web users ignore advertisements that appear in their visual field of attention. In 1998, Benway noticed that users who interact with online websites pay little to no attention to graphically salient messages. In her experiment, Benway (1998) asked six participants to search for specific information, such as an email address for a hotel. The target information was distributed between textual hyperlinks and banner ads, such as the one shown in Figure 5.



Figure 5. An example of stimuli used by Benway (1998).

Her results showed that participants located the target information in 94% of the trials when it was presented as hypertext, but with a success rate of only 58% when presented as a banner ad. Furthermore, her participants rated the search to be more difficult when the target information was presented in a banner ad. The Benway (1998) study results seemed to indicate that users ignored messages

(usually advertisements) that were—ironically—deliberately designed to be salient and attention-getting. As users continued to ignore banner ads, advertisers pursued more aggressive techniques to attract users; banner ads began employing animation to take advantage of users' attention capture reflexes caused by movement or an abrupt appearance of stimuli (Hillstrom & Yantis, 1994; Jonides & Yantis, 1988; Titchener, 1908). However, a number of studies found that animation did not affect ad avoidance nor increase brand recall (Bayles, 2002; Burke, Hornof, Nilsen, & Gorman, 2005; Spool, Scanlon, Schroeder, Snyder, & DeAngelo, 1997).

For example, Burke et al. (2005) asked their participants to select from a list of news headlines the one headline that most corresponded to a cue statement. In addition, they embedded two banner ads that were either static, animated, or neutral gray rectangular (see Figure 6).



Figure 6. An example of the search task used by Burke et al. (2005). The sentence on the top was used as a precue that disappeared immediately as the list of headlines appeared.

Their results showed that search time during the static and animated banner conditions were significantly longer when compared to the gray banners, but with no difference between static and animated. Using an ad memory post-test, they found that the recognition rates for static banners were actually higher than animated banners. Using eye-tracking analysis¹ Burke et al. (2005) found that the number of eye fixations on gray rectangles did not differ from the static nor animated banner. Combined with the search time results, which showed a significant increase in search time for the ad banners (static and animated) compared with the gray rectangles, the authors concluded that the participants processed the content of the ad banners *peripherally* rather than looking at them directly. Therefore, in this study, while banner ads affected the participants' response to the web content, animation did not increase their response towards the ad; rather it had an adverse effect on brand awareness as compared with the static banner ads.

Although it is clear that ad avoidance is a real and replicable phenomenon, it is not clear what *causes* it to occur. The next section will review potential explanations that have been reported in the literature.

Ad Avoidance as a Function of Repetition

There has been an ongoing debate in the literature about the effect of stimulus repetition (primarily of a persuasive nature) on the users' response to that

¹ Burke et al. (2005) performed gaze analysis *manually* by playing back screen capture videos of the participants' trials while monitoring the superimposed crosshair that represented the participants' gaze. Since the threshold for a fixation is between 200 milliseconds (ms) to 500 ms, it is unclear whether they could compute this data accurately or reliably.

stimulus. One model, named here as the "asymptotic habituation model", argues that repetitive stimulus exposure decreases the likelihood that one will direct their attention to that stimulus as a result of habituation, which is defined as "... response decrement occurring as a result of repeated or continuous stimulation..." (Horn, Hinde, & Centre, 1970, p. 3) (see Figure 7). Since the perceived information is already represented in memory storage, the stimulus is no longer salient, which in turn prevents the user from orienting their response to that stimulus (Sokolov, 1963).

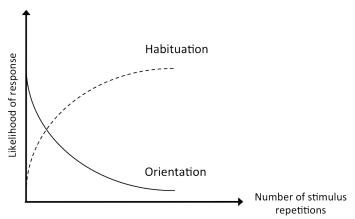


Figure 7. Depiction of the asymptotic habituation model. Stimulus repetition leads to lower likelihood of response to that stimulus as a result of habituation.

Studies in the ergonomics domain investigated how habituation affects perception and compliance to hazard labels. One of the earliest studies in this field found that users gradually became less compliant with a warning signal over time, thus demonstrating habituation (Thorley, Hellier, & Edworthy, 2001). In their experiment, the researchers posted a warning sign on one of two doors, asking the students to use the other door instead. Over time, the students began to ignore the warning sign and used the "defected" door to exit the classroom. Contrary to the researchers' initial hypothesis, changing the sign's appearance did not reverse the

habituation effect as the students' compliance rate continued to decrease. The authors argued that social influence might have confounded their results.

In a follow-up study (Thorley, Hellier, Edworthy, & Stephenson, 2002), the authors presented the participants with a set of pictorial and text warnings followed by another set with different warnings. Using galvanic skin response measurement, the authors were able to record the habituation progress expressed by a decrease in orientating response. When a new set of stimuli was presented, the results showed a significant dishabituation effect, represented via an increase in the participants' orienting response. However, the response in the following trials (using the new stimuli) was sharply reduced as compared to the pre-testing condition.

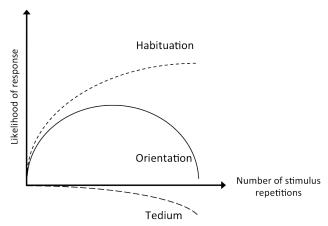
These results coincide with other studies that found a change in the warning appearance may disrupt the habituation process and increase the likelihood that the users will respond accordingly (Wogalter et al., 1995; Wogalter & Brelsford, 1994; Wogalter, DeJoy, & Laughery, 1999).

A similar pattern of response was found in the marketing domain as well (Appel, 1971; Buchanan & Morrison, 1988; Burke & Edell, 1986; Chatterjee, Hoffman, & Novak, 2003; Fry, 1996). A study by Buchanan & Morrison (1988) on response to direct mail found that following the first round of mailing, the response to a second mailing decreased by 50%, and continued to decrease with each subsequent round. However, Buchanan & Morrison (1988) explained this response pattern by a process of self-selection: people who are likely to respond to an offer will do so in the first round of mailing. Hence, subsequent rounds target people who are less likely to respond, which explains the high falloff rates.

Chatterjee et al. (2003) found the same pattern of response in reaction to banner ads. Chatterjee et al. (2003) had a rare access to a commercial website (the website's name is undisclosed) for 7.5 months, while monitoring the CTR of nearly 22,000 registered users on two commercial banner ads (also undisclosed). The authors found that repetition had a negative effect on CTR up until the 10th exposure, after which the CTR reached a plateau.

However, these finding stand in contrast to the results obtained by Drèze & Hussherr (2003) who found that repetition of banner ads increased their participants' recall rate of the ads and the brand awareness. These finding support the habituation-tedium theory, where habituation plays a *positive* role in the users' response.

The habituation-tedium theory (Berlyne, 1970; Rethans, Swasy, & Marks, 1986; Sawyer, 1981) argues that a response to a new ad is based on an interaction between habituation and tedium. Novel stimuli will result in a low orienting response because they instigate uncertainty and tension. Habituation serves a positive role in reducing the negative attitudes, resulting in more liking of the stimuli, which in turn leads to an increase in user response to the stimuli. However, with increased repetitions, boredom increases (tedium), which leads to a decrease in liking and orienting response (see Figure 8).



 $\label{eq:Figure 8.} \textbf{A modified illustration of the Habituation-Tedium Theory.}$

According to this theory, habituation has the most impact in early stages of exposure and tedium is most powerful in later stages. The result is an inverse U-shape form of response to stimulus as a function of liking.

The debate in the marketing literature about the effect of ad repetition has existed for many years. In a review article by Pechmann & Stewart (1988), the authors noted that the source of the debate arose from the use of different dependent variables, whether "attention, immediate recall, delayed recall, cognitive response, immediate brand attitudes, delayed brand attitudes or sales" (Pechmann & Stewart, 1988, p. 286). Indeed, while Chatterjee et al. (2003) used CTR as their primary dependent variable that indicated an immediate, short-term effect for banner ads, Drèze & Hussherr (2003) used recall measures that showed an effect in the long-term interaction between the users and the brand. However, both of these studies used real brand names as their stimuli (in the case of Chatterjee et al., 2003, the brands are not known), which may have impacted the internal validity of the results since participants had already been exposed to the brand names prior to the experimental manipulation. In addition, there have been no recorded attempts to

combine behavioral and cognitive dependent measures to assess the effect of online banner ad repetition on users' response.

The Effect of Search Type on Ad Avoidance

Because a significant portion of online advertising depends on users' search activities, it is important to investigate what effect (if any) different search tasks have on ad avoidance.

Marchionini (2006) defined three types of search activity: Lookup, Learn, and Investigate (see Figure 9).

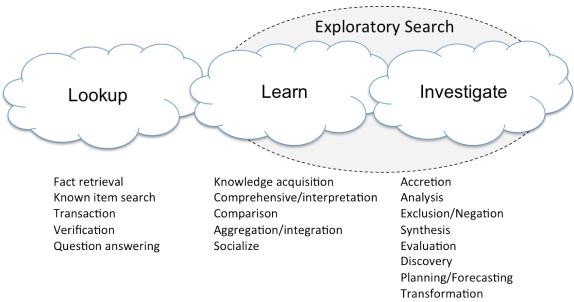


Figure 9. Model of search activities (Marchionini, 2006).

Lookup search queries return "discrete and well-structured objects such as numbers, names, short statements, or specific files of text or other media" (p. 42) and are often referred to as "known-item" search. Hence, a known-item search is defined here as a single query that returns a discrete and well-defined answer. An example of a known-item search query is "what is the capitol of Moldova?" or "how many people live in Bhutan?"

However, learning search requires the user to perform multiple iterations of their queries and to interpret a variety of information representations while performing qualitative judgments of the information retrieved. A learning search task would be "Is there a good Ethiopian restaurant in Washington DC?" While investigative search tasks are similar to learning tasks in that both require multiple iterations and qualitative judgments, investigative search tasks are often conducted over a long period of time to support long-term planning or identify gaps in information. An investigative search task would be, "Given current economic conditions, should I purchase a new house in Chapel Hill?"

According to Marchionini's model of search activities, Learn and Investigate search tasks yield exploratory search behavior because they require making a series of queries in combination with browsing strategies. Therefore, an exploratory search is defined as a series of investigative queries involving a particular topic. As will be discussed later, the qualitative difference between known-item and exploratory search could affect the orienting response to banner ads and advertising in general.

Based on field observations, Spool et al. (1997) speculated that "surfing" was more likely to yield a high click-through rate than "information retrieval" activity (p. 91). However, they did not clarify why this was likely to happen nor put forward any data to back this claim. A year later, Janiszewski (1998) examined whether different search behavior affected the product perception in print shopping catalogs.

Specifically, he compared known-item search versus exploratory search.

Janiszewski (1998) postulated that known-item search was driven by top-down

attention while exploratory search was stimulus driven via salience² (a review of top-down and bottom-up attention theory will be presented in the next section). Consequently, the information layout may affect known-item searchers differently from exploratory searchers. Janiszewski (1998) found that, when engaged in a known-item search, the amount of time spent by his participants was a function of the saliency of the targeted item and its relevance to search task (higher saliency led to less search time). However, when engaged in an exploratory task, the participants' attention (as measured by amount of viewing time at the display) was a function of the number of items on the display and their proximity to one another.

Similarly, Pagendarm & Schaumburg (2001) examined whether known-item searchers are less likely to notice online banner ads than exploratory browsers. They asked the participants in the known-item group to answer as many questions as possible within 15 minutes. In contrast, participants in the exploratory group were asked to freely browse through the web pages and read the content pertinent to their personal interests. Both groups were exposed to the same website (see Figure 10). The top section of the webpage remained fixed and showed a single animated banner ad for 45 seconds, and each banner ad was presented only once during the experiment (a total of 16 banner ads were presented). The bottom section of the webpage showed the content area.

_

² Salience is a psychophysical measure of a signal-to-noise ratio (SNR) between an object and its surrounding. The higher the SNR, the more salient is the object, thus making it easy to notice.



Figure 10. A screen shot of the stimuli used by Pagendarm & Schaumburg (2001).

Using a free recall and recognition test, Pagendarm & Schaumburg (2001) found that the participants in the exploratory condition had significantly better recall rate then participants in the known-item condition. Specifically, with the exception of text slogans, participants in the exploratory condition scored better in recalling the names of companies, products, services, colors, and pictorial motifs. In addition, on average, participants in the exploratory condition scored higher banner ad recognition rates.

However, the study had several methodological issues. The recognition test did not have any distracters (i.e., banner ads that did not appear during the experiment), and since the participants were asked to rate their confidence in seeing a banner ad previously, the authors were confounding between confidence and recognition. In addition, as the authors reported, participants in the known-item group spent considerably more time looking at the paper sheet to review the search questions. Since the banner ads were rotated automatically, the known-item group

of participants had substantially less time to look at the individual banner ads than did the participants in the exploratory condition.

Danaher & Mullarkey (2003) used a within-subject design to compare the effects of search type on advertising memory. For this purpose they constructed a mock student portal and included two banner ads that advertised mock brands, one for a computer manufacturer (part 1 of the study) and another for an online magazine (administered in part 2). In the first part of the experiment, the participants were asked to perform a series of known-item search tasks (no examples were reported), followed by a series of recall and recognition questionnaires. For the second part of the study, the participants were instructed to evaluate the website. Because the authors assumed that participants would be more likely to pay closer attention to the ads in the second part, the authors asked them to complete two general tasks that required a high mental workload (typing text they remembered viewing on the website along with their thoughts about the website).

The authors found a significant difference in banner ad memory between the known-item search and surfing (exploratory) conditions: participants in the exploratory condition had higher memory scores than did participants in the known-item search condition. However, since Danaher & Mullarkey (2003) did not counter balance the order of the conditions, a training effect cannot be ruled out, which may have confounded their results.

When examining the eye gaze pattern of web page visual elements, Yesilada, Jay, Stevens, & Harper (2008) found a similar pattern when comparing known-item and browsing behaviors. Participants in both conditions spent less time looking at

the web page's peripheral sections than the main content section. However, in the browsing condition, participants had a significantly higher fixation count in the peripheral section of the web page as compared with the known-item condition. However, the task that was given in the exploratory browsing condition was to evaluate the web page's layout, which essentially required that participants look at the web page's peripheral sections. Therefore, the effect that was found is a result of the specific instructions that were given to the participants rather than to changes in the navigation style.

Similar to the methodology used by Pagendarm & Schaumburg (2001),

Calisir & Karaali (2008) asked participants in the known-item condition to answer

10 multiple-choice questions and the participants in the exploratory condition were

asked to browse aimlessly. All the banner ads advertised a single Turkish

commercial bank and had three variations (see Figure 11).



Figure 11. The website used by Calisir & Karaali (2008) and the three banner ad variations.

To measure the orienting response to the ads, Calisir & Karaali (2008) used a recognition test where eight banners were presented but only *one* of which was included in the experiment. The participants were asked to indicate whether they saw the ad and their degree of confidence. Later their responses were combined to create a six-point scale ranging from 1 (very sure that the answer was correct, however, the answer turned out to be incorrect) to 6 (very sure that the answer was correct, and the answer turned out to be correct).

Unlike the previous findings (Danaher & Mullarkey, 2003; Pagendarm & Schaumburg, 2001; Yesilada et al., 2008), Calisir & Karaali (2008) reported that participants in the known-item condition had higher recognition scores than did the aimlessly browsing participants, but only for the banner ad that included URL and service information (the bottom of the three banner ads shown in Figure 11). It is possible that the inclusion of additional information in the banner helped the known-item searchers in their task, but since the authors did not include any examples of their search tasks this remains a speculation.

Finally, the most recent publication on the topic of known-item versus exploratory search was of an eye-tracking study conducted by Owens, Chaparro, & Palmer (2011), who examined the phenomenon of ad blindness in text ads. Owens et al. (2011) used a within-subject design and asked their participants to perform a mix of "exact" (known-item) and "semantic" (exploratory) searches. An example for a known-item search was "Your friend is planning a trip to the Island of Kauai and is searching for activities to do on the island. Find the 1080 page guide for Kauai." An example for a semantic search was "You and a friend are traveling to Maui and wish

to visit Haleakala National Park. Find where information about the cost, when the park is open to the public, etc. could be located." The answers for these searches were equally distributed between the content areas of the web page and the text ads.

To measure the effect of their manipulation, the authors used a combination of post-test interviews and eye gaze data. In the post-test interview the participants were asked to describe the web page layout (i.e., content, text ads, menu) rather than specific information about the ads themselves. Their results showed no significant difference in eye fixations on the ads area between exact and semantic search conditions. Unfortunately, the authors did not report the results from the post-test interview, which could further clarify results of their investigation. One possible explanation for the lack of significant difference in eye fixations is that the semantic searches were not exploratory in nature but rather a collection of multiple known-item queries. Using the semantic search task example, the participants were asked to find the cost of the trip, hours of operation, and its location—all queries that led to discrete and well-structured answers. Therefore, instead of manipulating the search behavior, the authors manipulated the task *duration*, as evidenced by their reported results of a significant difference between exact and semantic search where exact search tasks took less time to complete.

In summary, while the majority of the studies reported here indicated that participants who engaged in a known-item search task were less likely to notice advertisements online than were participants engaged in an exploratory search task, the results reported here are inconclusive (see Table 1). Furthermore, as

reported earlier in this section, several of the findings are questionable due to issues with the studies' internal validity.

Table 1

Summary of the investigation of the difference in orienting response to banner ads between known-item and exploratory search tasks.

Publication	Exploratory > Known-item	Known-item > Exploratory	No difference
Pagendarm & Schaumburg (2001)	X		
Danaher & Mullarkey (2003)	X		
Yesilada, Jay, Stevens, & Harper (2008)	X		
Calisir & Karaali (2008)		X	
Owens, Chaparro, & Palmer (2011)			X

A goal of this dissertation is to clarify whether there are any differences in ad perception between known-item and exploratory search tasks, while ensuring a proper experimental design to avoid the previously reported mishaps.

The Role of Salience in Orienting Response to Banner Ads

It has long been the assumption that bigger, brighter, and more interactive ads will attract the attention of potential consumers to banner ads. That notion has roots in a plethora of research conducted on traditional print advertising (e.g., Barlow & Wogalter, 1993; Berdie, 1992; Lohse, 1997; Lutz & Lutz, 1977; Valiente, 1973). However, based on the rapid decline in CTR and on the evidence of the studies presented here, online users seem to become immune to the use of bold graphical elements in online banner ads. While it is not clear how habituation affects the orienting response to banners ads over time, there is evidence that online users quickly adapt to new forms of advertising by blocking them from their focal point. The result is a race to the bottom, where ads become more and more invasive and aggressive as advertisers scramble to attract the users' attention.

In a series of experiments, Folk, Remington, & Johnston, (1992) found that participants automatically and efficiently ignored highly salient objects that were deemed to be irrelevant to the main task. In fact, a few studies found that the existence of highly distracting elements *improved* the participants' visual search performance (Pashler, 2001; Van Orden, Divita, & Shim, 1993).

For example, Van Orden et al. (1993) examined whether adding flashing and luminance signals to targets that are already color-coded would improve the participants' search performance or rather lead to a deterioration of performance. The participants were asked to locate the quadrant of the plot with the greatest number of red, blue, or yellow hat-shaped symbols (see Figure 12). In some conditions, a subset of the symbols was highlighted or began flashing while other symbols remained static.

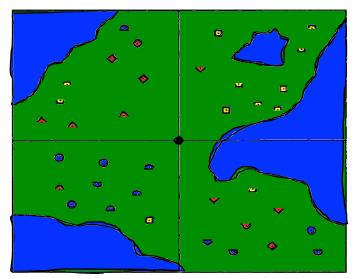


Figure 12. Geographical map used in the experiment by Van Orden et al. (1993).

As expected, highlighting and flashing increased the search efficiency as evident by the lower search time. However, counter intuitively, flashing and highlighted targets did not deteriorate the search performance when the

participants were asked to search for non-flashing or non-highlighted targets. The participants reported that they utilized this apparent distraction to *guide* them to the relevant, non-flashing targets.

In this case, it may be argued that habituation was taking place because a representation of the object was already available in the participants' memory storage (whether long or short-term). As a result, that object would be less likely to activate bottom-up attention due to its weak salience. Instead, users might have utilized these remembered objects via a top-down mechanism to attend only to information that did *not* flash or highlight. From this line of reasoning, one could speculate that the more flashy and colorful the ad is, then—ironically—the less likely it will attract the users' attention.

Indeed, a recent study by Tullis, Siegel, & Sun (2009) lends support for this argument. In a series of experiments, Tullis et al. (2009) asked their participants to find specific information embedded in a financial web portal. Half of the participants were shown a message that included a face and the other half were shown a message that did not include a face (see Figure 13).

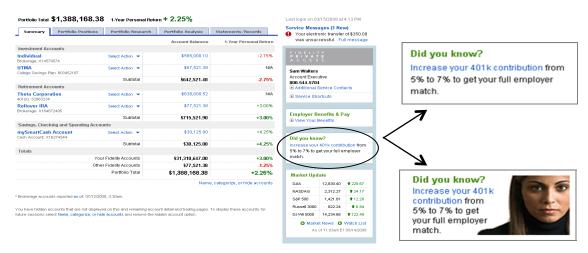


Figure 13. An example of the stimuli used at Tullis et al. (2009).

Their results showed a remarkable difference between the face condition and no image condition in terms of accuracy (78% vs. 93% correct, respectively) and time-to-completion (54 sec. vs. 37 sec.). Therefore, the use of a facial image in the banner decrease its effectiveness and the user experience compared with text messages.

Based on these results and the previous findings from the exploratory vs. known-item search literature, I postulate that users who engage in known-item search are more likely to notice low salient ads than highly salient ads. The opposite is true for users who engage in exploratory search; in this case, highly salient ads are more likely to capture the users' attention than low salient ads. An exhaustive search of the literature did not yield any published attempts to explore this interaction.

This section provided an overview to the problem domain of ad avoidance. Despite the decline in CTR, attempts to reverse this trend have met with limited success. More studies have revealed banner ads' lack of effect on users' attention than positional solutions that could mitigate the situation. Even when positive effects were found they were often contaminated by confounds. Consequently, after 15 years of research into the problem of online ad avoidance, there is little that academic researchers can offer to our colleagues in industry. This dissertation attempts to change things.

To do this, I focused my investigation on three core factors that affect the users' orienting response to banner ads: habituation (as a function of ad repetition),

engagement in a known-item search vs. exploratory search, and salience features of banner ads.

The following section delves deeper into the cognitive psychology literature to provide a common platform for discussing concepts such as attention and mental workload, among others, and also provides further explanations the ad avoidance mechanism.

Theories of Attention and Mental Workload

The general consensus of the published research seems to indicate that ad avoidance results from "top-down [attention] processes" (Pagendarm & Schaumburg, 2001) or "information overload" (Müller et al., 2009). Yet these labels do not provide a full description of the cognitive processes that take place during an advertisement avoidance behavior.

Therefore, the primary goal of this section is to bridge this gap and provide a theoretical foundation to my research endeavors. In addition, I also review the historical development of attention theories and their application in human-computer interaction research.

Fundamentals

This section reviews the fundamentals of visual attention research. The modern cognitive research literature discussed here provides theories and concepts that could be productively applied to investigate the phenomenon of ad avoidance.

What is Attention?

Despite its common usage in both popular and academic literature, there is little agreement among academic scholars as to the precise definition of "attention" (Styles, 2006). One of the earliest definitions was by the founder of American psychology, William James (1890):

"Every one knows what attention is. It is taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneous possible objects or train of thought. Focalization, concentration, of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others, and is a condition which has a real opposite in the confused, dazed, scatter-brained state which in French is called *distraction*, and *Zerstreutheit* in German" (James, 1890, pp. 403-404).

This definition implies several interesting assumptions. Firstly, it suggests that the attention mechanism acts as a *selective filter* of our surrounding environment. Secondly, this process requires *active* and *conscious* control. Thirdly, this definition suggests that there is a *limited capacity* to our mental resources. While some of James' insights were proved to be scientifically valid (as will be discussed later in this paper), others were not. For example, as early as 1908, Titchener (1908) discussed the *automaticity* of attention processes as a result of a novice object in a visual field or sudden movement. As to the assertion of a limited capacity and selectivity—that would depend on the *type* of attention we are referring to. For example, we could easily look at an object while *simultaneously* listening to a conversation behind us.

Consequently, for the purpose of unity, throughout this paper I will use the term *attention* as defined by Jenkin & Harris (2001):

"Attention implies allocation [of] resources, perceptual or cognitive, to some things at the expense of not allocating them to something else" (Jenkin & Harris, 2001, p. 1).

As described earlier, there are several types of attention. These include selective, divided, sustained, alertness, and more. The following section will focus on selective and divided attention, as these are most pertinent to the problem domain of ad avoidance (a review of other types of attention is available at Jenkin & Harris, 2001; Styles, 2006; Wickens & McCarley, 2008)

Selective vs. Divided Attention

The core difference between these two types of attention is that selective attention refers to the ability to focus on a *single* stimulus while ignoring others and divided attention refers to the ability to process information from multiple channels *simultaneously*. Historically, studies in selective attention focused on the auditory channel rather than the visual channel. The reason is that, unlike the visual channel, where we can easily shift our eyes from one focal point to another or shut them, our ears are fixed in one place. Therefore, auditory attention must rely on neural mechanisms rather than mechanical (Broadbent, 1971).

Early studies in attention found that different voice pitch and physical location of the message source were utilized to select which auditory information to listen to (Broadbent, 1952; Broadbent, 1954; Cherry, 1953). These findings led to one of the earliest theories of attention: the filter theory, also known as early selection theory (Broadbent, 1958). Inspired by information theory (Shannon & Weaver, 1949)³, the main idea behind the filter theory is that stimuli are first being processed by our senses using dedicated sensors that attenuates to the physical

-

³ Information theory uses mathematical modeling to quantify information. In addition, it describes the communication flow from a source to a destination while accounting for the channel capacity, signal-to-noise ratio, and other factors that may affect the quality of data being transmitted.

property of the information (pitch, loudness, and spatial positioning). The short-term memory stores this information that is later passed to the "filter". The filter selects *one* input line and allows it a direct access to the limited capacity channel. Information (input line) that was not selected is retained in the short-term memory for a few seconds, during which time degradation begins taking place. After being processed by the limited capacity channel, the input line can be associated with information stored in the long-term memory, which eventually leads to an output in the form of muscular activation (see Figure 14).

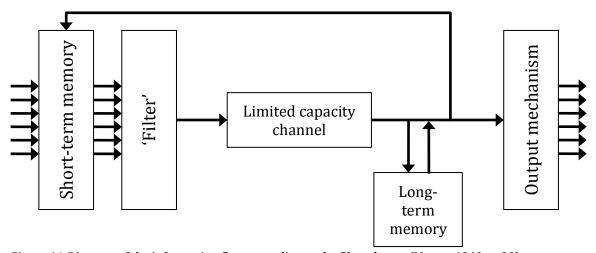


Figure 14. Diagram of the information flow according to the filter theory (Moray, 1969, p. 29).

Findings that supported this theory included a series of experiments conducted by Cherry (1953), who asked his participants to listen to two different auditory messages presented to each ear simultaneously. In addition, he instructed the participants to repeat one of the messages out loud without delay (shadowing task). His results showed that people were able to perform this task fairly well. In addition, he found that when asked to describe the content of the rejected message, the participants were not able to do so, except that they had noticed a sound present.

However, one of Cherry's (1953) findings contradicted the filter theory's main hypothesis: the participants noticed when the rejected message included changes in gender or tone. That result suggested that a *parallel* processing of multiple stimuli can take place beyond the filter stage—in other words, divided attention.

In a follow-up study by Moray (1959), participants were presented with a list of words to each ear pronounced by the same voice. Again, their task was to repeat the words heard in one ear as soon as they could while ignoring the words heard in the second ear. Although some of the words were repeated up to 35 times, the participants were not able to identify the words from the rejected channel but performed remarkably well when asked to identify words presented in the attended auditory channel. However, when the rejected channel included the participants' name, they were able to report that accurately. In a study by Loftus (1974), participants were presented with the instruction "You may stop now" in the rejected channel; participants complied only 6% of the time compared with 33% when presented with the same instruction that include the participants' name.

Further evidence that people can process some semantic information in parallel to a selective attention task came from the study by MacKay (1973; cited in Pashler, 1999). MacKay presented his participants with ambiguous sentences in the attended channel and potentially disambiguating information in the rejected channel. His results showed that information presented in the rejected channel biased the participants' interpretation of the ambiguous sentences presented in the attended channel. For example (taken from Pashler, 1999, p. 52), subjects tended to

interpret the sentence "They threw stones toward the bank yesterday" as being about a financial institution when the word "money" was in the unattended channel, and as being about riverbanks when the unattended word was "river". However, this effect was found only when a single word was presented in the rejected channel and disappeared when a series of words were presented (Newstead & Dennis, 1979).

Despite this evidence of parallel processing, it is clear that the capacity to process information from the unattended channel is strikingly lower than when processing information from the attended channel. In response to the earlier findings, Treisman (1964a; 1964b; 1964c) modified the filter model to include an attenuated filter that weakens the strength of some signals rather than blocking those that do not meet the criteria altogether (see Figure 15). If the signal was not completely blocked, it might have sufficient strength to activate the representation of certain words and pass the threshold of consciousness (e.g., the participant's name).

The work of Triesman and others (Deutsch & Deutsch, 1963; Reynolds, 1964) eventually led Broadbent to modify his theory by incorporating Triesman's ideas about the attenuating filter such that the filtering process represents "evidence" for the stimulus. Although the evidence may suggest a particular state and form of the stimuli is more likely than others, this is not a deterministic process.

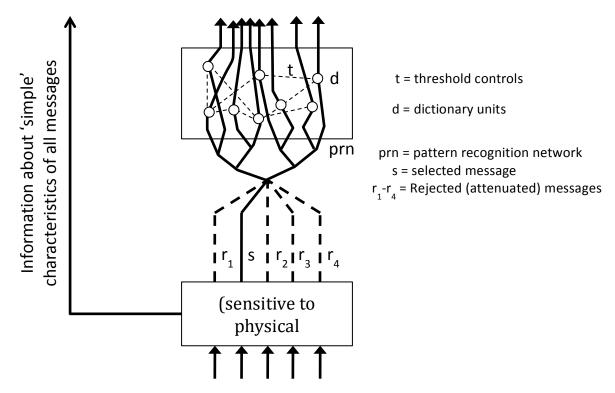


Figure 15. Treisman's input selection model (copied from Moray, 1969, p. 31).

In addition, Broadbent expanded the role of the selective filter stage to include two selection mechanisms: pigeonholing and categorizing (Broadbent, 1971; Broadbent, 1982). Pigeonholing referred to the automatic assignment of unrelated input to the incoming information and categorization referred to the expansion or modification of existing classifications to help interpret new information.

The final output of this information flow resulted from what Broadbent referred to as "category state", which was the final representation that was triggered as a result of pigeonholing and/or categorization. Consequently, the category state could be biased and result in an erroneous output. For example (taken from Styles, 2006, p. 38), if a participant was asked to listen for the name of an animal, then they would very rapidly respond to animal words irrespective of the

voice in which the words were spoken and they would be more responsive to animal words than words associated with building supply.

However, the view that a restriction of the limited capacity channel to serially process one input at a time eventually gave way to increasing evidence that parallel processing may be occurring at higher cognitive operations than was previously suggested by the filter theory. In the next section, I will examine models of visual search that explain how people search for information in our surrounding environment and attend to specific objects.

One of the most common daily activities that human engage in is visual search, whether it is looking for car keys on a table or a knife in a luggage X-ray screening. Simply put, visual search is "...an effort to detect or locate a target item whose presence or position within the search field is not known a priori" (Wickens & McCarley, 2008, p. 63). The filter model attempted to explain visual environment processing by analyzing the environment's individual components and separating out the various object attributes. However, there was little explanation as to how the various object attributes combine to create an image of a unified world.

Feature integration model

Among the most important and well-cited models of attention (over 5000 citations) is the feature integration model (see Figure 16) proposed by Treisman & Gelade (1980).

The feature integration model postulates that sensory features, such as color and orientation, are processed in parallel and pre-attentively. In this case, a target

appears to "pop out" with little or no conscious effort (for example, when searching for a red object surrounded by green objects).

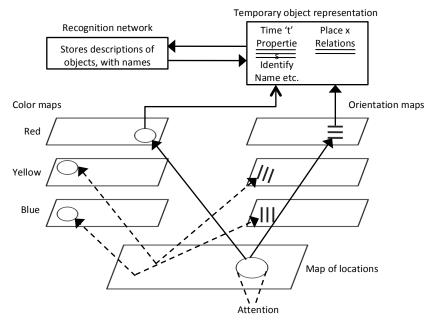


Figure 16. Feature integration model (copied from Treisman & Gormican, 1988, p. 17).

However, in cases when one needs to locate a target that shares multiple features with its surrounding (for example, a red square surrounded by red circles and green squares), these features must be combined into a conjunction. This can be done either by matching the objects' features to a predicted object based on already stored descriptions of the object or by focusing the attention in a particular area within the "map of locations", which represents the location of all the features.

Focusing on one location of this map leads to retrieving a certain set of features used to create a "temporary object representation", which is matched to the content of the object file. Finally, features can conjoin independently, without the use of attention. In this case, the conjunction is more likely to lead to erroneously

interpreting the surrounding objects, which Treisman and Schmidt (1982) refer to as "illusory conjunction".

To test their theory, Treisman & Gelade (1980) ran a series of experiments,

using reaction time as the dependent variable. They showed that while the reaction time for conjunction search tasks (searching for a green "T" among green "X" and brown "T") increased linearly with the number of distractors, the reaction time for search disjunction tasks (looking for a blue "S" among green "X" and brown "T") was independent of the

Conjunction

Conjunction

Disjunction

pos

neg

neg

pos

neg

pos

neg

Display size

number of distractors (see Figure 17). In addition, they found that the absence of a target (annotated as

Figure 17. Search times as a function of type of task and display size (Treisman & Gelade, 1980, p. 104).

"neg" in Figure 17) led to an increase in reaction time, but again the effect was much stronger for conjunction tasks as compared with disjunction tasks.

The feature integration model quickly spawned a vast number of follow-up studies and became a milestone in cognitive psychology (see Quinlan, 2003 for a complete review). A number of models emerged, all based on the premise that the visual surrounding is first encoded into a set of basic perceptual features. This led to the idea that visual search can be systematically *guided* by the properties of the surrounding object and the type of task that the user is engaged in.

Guided Search 2.0

The Guided Search 2.0 model (GS2) (Wolfe, 1994), is one of the leading models that explains guided search while incorporating many of the concepts first

suggested by the filter and feature integration theories. The GS2 is divided into three stages: Input Channels, Feature Maps, and Activation Maps (see Figure 18).

Input Channels: Tuned input channels filter each feature, whether by color, orientation, size, etc. Input channels absorb information from the stimuli according to different categories. For instance, in the case of orientation perception, different channels respond to different angles. If a line is within an angle orientation range of -45° to 45°, the line will be categorized as "Steep".

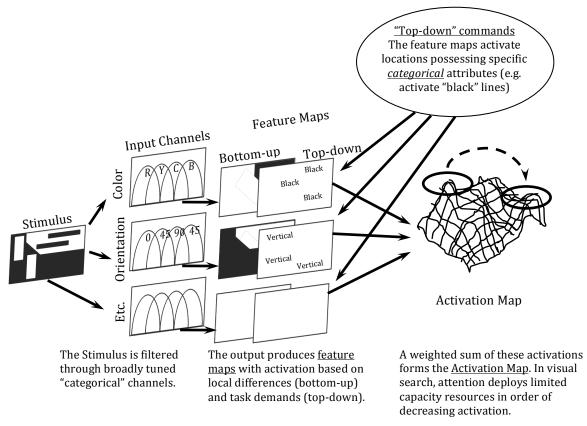


Figure 18. Basic components of the guided search 2.0 model (copied from Wolfe, 1994, p. 205)

<u>Feature Maps</u>: As originally suggested by Treisman (1980), feature maps are independent representations of basic visual features, such as color, spatial orientation, etc. Each component within the feature map creates some level of activation that in turn attracts attention to that unique feature. The greater the

activation, the more likely it is that attention will be directed to that component. Wolfe (1994) describes two components of activation: bottom-up and top-down.

Bottom-up activation measures the activation caused by a distinctive item compared to its surroundings, without the need for any previous knowledge about that item. For example, a single green square among a group of red squares will generate a larger activation than the surrounding red squares, which will lead to its detection during a visual search task. For a difference in activation to be detected, the item must overcome a certain difference threshold, which Wolfe (1994) referred to as *preattentive just noticeable difference* (Pjnd). Pjnd's in parallel search tasks are greater than the perceived jnd in a standard psychophysical task (Weber, 1846). For example, although humans can detect an angular difference of less than 1°, the Pjnd for parallel search is in the range of 10°-15°. In other words, when humans are given a task to judge whether there is an angular difference between two items, they are capable of differentiating between items that have less than 1° difference. However, when given an array of homogenous items, the target must have an angular difference of 10°-15° to attract attention.

Bottom-up activation is effective in a homogenous environment, where any irregular item simply "pops up". However, in a heterogeneous environment, bottom-up activation may be considered "noise". For example, when trying to locate a green square among a variety of colored squares, bottom-up activation would make it very difficult to locate the item, simply because all items generate an equal amount of activation that cannot easily be categorized. To "answer" this problem, a second activation mechanism is used: top-down activation.

According to Wolfe (1994), top-down activation is achieved by tuning one of the input channels to a particular attribute that best *differentiates* the target from its distracters, unlike bottom-up activation, which simply depicts the target with the *largest* activation. For example, in a visual search for a 20° rightward tilted line among vertical line distractors, the 20° line will be categorized by the visual channels as "steep" and "right". However, a perfectly vertical line (0°) will produce a larger "steep" response, but not "right". Therefore, the correct search behavior will be to monitor lines with the "right" property.

Activation Map: The origins of the activation map mechanism could be traced back to Deutsch & Deutsch (1963), who claimed that each stimulus perceived by the senses is integrated into a discriminatory mechanism. This mechanism becomes "excited" by the attribution of the perceived stimuli. The higher the activation, the more likely that stimuli will attract our attention. Similarly, Wolfe (1994) proposed that an activation map is created by the weighted sum of all the top-down and bottom-up components. As seen in Figure 18, activation maps look like a topographic map, where the hills are areas of high activation that attract attention. The activation map does not contain any information regarding its source, such that an area in the map that derives its source from color looks the same as activation caused by orientation. In a parallel search, the attention will be attracted to the highest level of activation, regardless of the number of items. Thus, parallel search reaction time is not a function of the number of distractors. In serial searches, items that have multidimensional features, such as color and size, would have higher activation than a singleton. Therefore, items are examined in order of their

activation levels, from high to low, which can explain an increased reaction in a serial search task as a function of set size.

While theories of attention and visual search can explain how we shift our attention from one object to another, as first suggested by (Cherry, 1953), voluntary and involuntary shifts of attention are heavily influenced by our limited mental capacity. The next section will review some modern mental workload theories and their application to the problem domain of ad avoidance.

Mental Workload

As discussed previously, people have a limited capacity to process concurrent sources of information. Therefore, the more activities they engage in, the less effective they become (Simon, 1967). Every time they shift attention from their primary task (e.g., due to interruption), it takes time to readjust their focus. Therefore, when most users turn their attention to a webpage's content, they will most likely allocate their mental resources solely to that task.

Multiple Resource Theory

The multiple resource theory (MRT) suggests that humans have several capacity channels (Wickens, 1984; Wickens, 2002). Each channel is dedicated to a different sensation, such as vision or hearing. The main purpose of MRT is to predict the performance of people engaged in two time-shared and continuous tasks. MRT proposes four categorical and dichotomous dimensions that affect human performance, with each dimension divided into two discrete levels (see Figure 19):

stages (cognitive vs. response), perceptual modalities⁴ (auditory vs. visual), visual channels (focal vs. ambient), and processing codes (visual vs. spatial).

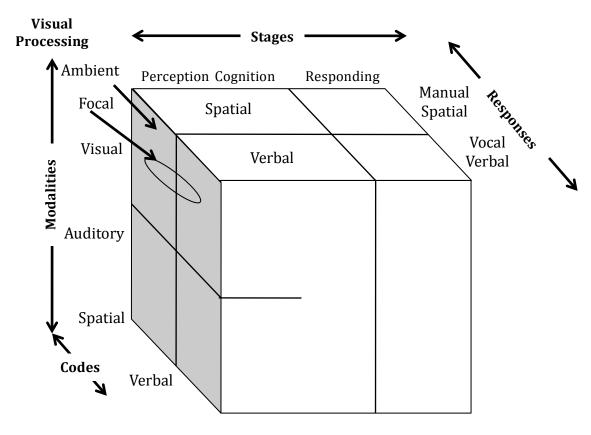


Figure 19. The structure of multiple resource theory (copied from Wickens, 2002, p. 163).

1. Stages: MRT states that the mental resources needed for perception and cognition are the same. These resources are functionally separate from those required for responding. For example, adding the requirement for a control room operator to vocally acknowledge the plant's state should not affect his ability to monitor and maintain the plant's operation.

39

⁴ It is important to keep in mind that these modalities are not completely separate from each other. In his 1984 paper (p. 87), Wickens cited a few studies that reported cross-modalities interferences, which supports the notion of a finite unitary attention capacity.

- 2. Perceptual Modalities: It is easier to divide one's attention between the eye and the ear as compared to dividing one's attention between two auditory channels or two visual channels. For example, driving while texting (two visual tasks) will more likely lead to performance deterioration compared to driving while listening to music (visual and auditory tasks).
- 3. Visual Channels: Within the visual processing modality, focal and ambient vision appear to feed separate mental resources. Focal vision is required for processing high-resolution information. In contrast, ambient vision involves peripheral vision and is used to sense orientation and self-motion. For example, one can easily read a webpage while being aware of the location of the sentence in relation to the page.
- 4. Processing codes: Spatial and verbal processes, whether functioning in perception, cognition, or response modes, depend on separate mental resources. For example, driving a car while dialing a phone requires engaging in two spatial tasks that interfere with each other. Replacing the manual dialing with voice-activated dialing would decrease the interference with the driving task.

The MRT model has been heavily used to develop and design multimodal interfaces, where two or more input modes are used to interact with the system.

These modes could include touch, speech, hand gestures, gaze, and haptic feedback.

Multimodal interfaces have been shown to improve usability (Karat, Halverson,

Horn, & Karat, 1999; Lai & Vergo, 1997), accessibility of user interfaces (Vitense, Jacko, & Emery, 2002), and decreased mental workload (Oviatt, Coulston, & Lunsford, 2004). In addition, MRT was also used as a theoretical basis to explain the superiority of audio-visual advertisements compared with single modality ads using visual or auditory alone (Leigh, 1991; Rau & Chen, 2006; Smith & Buchholz, 1991).

MRT can explain why online users ignore banner ads: both the banner ad and webpage content compete for the same mental resource pool. Therefore, the users' attention focuses more on the webpage's content structures, which are usually located in the middle of the page, rather than on banner ads that are usually located in the webpage's peripheral areas. However, MRT is not sufficient to explain the differences found between known-item and exploratory search tasks (if such a difference actually exists). A new mental workload theory proposed by Lavie, Hirst, De Fockert, & Viding (2004) attempts to bridge this gap by explaining why people are distracted by irrelevant information and what conditions encourage distraction; this theory is discussed in the next section.

Load Theory of Attention and Cognitive Control

For decades there has been a growing debate among cognitive psychologists between the supporters of "early-selection" and "late-selection" attention mechanisms. Early-selection theories (Broadbent, 1958; Cherry, 1953; Treisman & Gelade, 1980) postulate that people have a limited processing capacity, which forces them to select only a few stimuli for cognitive processing. Therefore, stimuli that are not attended to cannot be perceived. For example, Simons & Chabris (1999) asked their participants to watch a 75-second video clip of two groups of students,

wearing white and black shirts, passing a basketball to one another. In addition, they asked them to count the number of passes by either the "white" team or the "black" team. During the video, a person wearing a gorilla suit walked to the center of the screen, stopped, pounded their chest, and walked away; the event took 5 seconds. The researchers found that 46% of their participants did not notice the gorilla (the video can be viewed at http://viscog.beckman.illinois.edu/flashmovie/15.php).

However, many studies that examined the effect of irrelevant interferences found that people cannot voluntarily block irrelevant stimuli from being processed (e.g. Logan, 1988; Posner & Snyder, 1975). These late-selection models (Deutsch & Deutsch, 1963; MacKay, 1973) proposed that one's perception has an *unlimited* capacity to automatically process all surrounding stimuli in parallel. Selecting which stimulus to respond to is based on its relevance to the task at hand. Representations of the unselected stimuli are stored in the short-term memory and quickly decay.

To bridge the gap between early- and late-selection theories, Lavie et al. (2004) proposed a new model, which involved two mechanisms—passive and active—in rejecting irrelevant stimuli from visual attention.

- 1. Perceptual Load (passive): In tasks that involve high perceptual load—meaning a high number of multiple competing stimuli that engage the user's full capacity—there will be no capacity left for the processing of distractors. This results in an early-selection.
- 2. Cognitive Load (active): When the perceptual load is low, an active mechanism operates to reject irrelevant stimuli. Cognitive load utilizes working memory to maintain the current priorities set by

the user. Therefore, in low perceptual demand, but with high cognitive load (i.e., the existence of large number of distractors), the working memory capacity quickly reaches its limit, which results in increased processing of distractors. This results in a late-selection. Figure 20 illustrates the interaction effect of distractor

susceptibility as a function of cognitive and perceptual load.

Although both known-item and exploratory search involve cognitive and perceptual load, users who engage in an exploratory search are subject primarily to cognitive load since their task requires evaluation and synthesis of the information. An exploratory

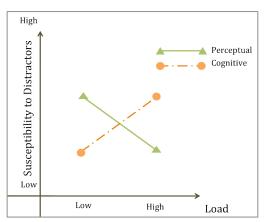


Figure 20. Interaction between different levels of distractors susceptibility as a function of low and high cognitive and perceptual load.

search example is "given the current financial situation, should I buy Nokia stock?" To answer this question, one would need to read about the status of the financial market, analyze the performance of Nokia, estimate their risk tolerance, and estimate how much money they are willing to invest. The information that was retrieved through those exploration steps is stored in the working memory; this could explain why exploratory search tasks are more likely to be associated with a high cognitive load.

In contrast, users who engage in a known-item search are primarily faced with perceptual load because they engage in a target-recognition activity. For example, if one is searching to find out who was the 42^{nd} president of the United

States, they would scan the text to find the number "42" and the word "president".

The surrounding text functions as masking stimuli, which leads to an increased perceptual load. Therefore, based on the load theory, known-item searchers are less likely to attend to banner ads than exploratory searchers.

However, contrary to the predictions of load theory, some of the studies that examined the difference between search type in online ad perception found a significant effect when the participants in the exploratory condition had a low cognitive *and* low perceptual load (Danaher & Mullarkey, 2003; Pagendarm & Schaumburg, 2001; Yesilada et al., 2008). In those studies, participants in the exploratory condition were asked either to aimlessly browse the website or evaluate the appearance of a website. Since those participants were not motivated to prioritize between relevant and non-relevant distractors, they were more likely to be drawn to the ads due to their saliency. But since the comparison was between known-item vs. exploratory search, aimless browsing could not be used as a task in this comparison because it did not involve any *search* activity. Thus, the comparison was made between completing a task and browsing aimlessly, rather than two distinct types of search behavior.

Moreover, it could be argued that true aimless browsing rarely exists, since most users are motivated by the desire to acquire knowledge as a result of their browsing activity. Whether they are interested to find today's weather forecast or catch up with the latest celebrity gossip, real-world users continuously prioritize the content to which they want to attend. Therefore, even if there is a difference in ad

perception between aimless browsers and known-item searchers, it is doubtful whether the effect has ecological validity.

In this section, I reviewed the development of attention theories from the dawn of the filter theory (Broadbent, 1958) to modern models such as load theory (Lavie et al., 2004). The reviewed literature in cognitive psychology provide the building blocks to explain the phenomenon of ad avoidance. While early theories were useful in explaining how users guide their attention to relevant stimuli (web content) and ignore irrelevant stimuli (banner ads), only recently have attention theorists been able to explain variability in ad avoidance results as a function of browsing behavior.

The next and final section reviews the use of physiological measures in human-computer interaction (HCI) research. Historically, investigating ad avoidance involved a combination of eye tracking and memory questionnaire. However, other physiological measures, such as galvanic skin response and heart rate, could also be used to measure the effect of banner ad exposure under a variety of experimental conditions. Therefore, the following section will examine these technologies and review how these measures were used to elucidate the problem domain of ad avoidance and in HCI research in general.

Physiological Measures in Human-Computer Interaction

While a variety of self-reporting and behavior logging techniques are prevalent in HCI research, there have been limited attempts to assess online consumer behavior using physiological measures such as galvanic skin response

and heart rate. These methods can detect an autonomous response within 15 to 75 milliseconds of exposure to a stimulus, a response that could otherwise be missed using traditional eye tracking measures (Lang, Potter, & Bolls, 2008). Also, there have been no published attempts to assess correlations among different biometric indicators of attention and habituation, such as eye movement, heart rate, and galvanic skin response. Integrating these methods may become essential to find small effect sizes in user experience research that continue to elude researchers.

Research in HCI and human factors is deeply rooted within the behavioral and cognitive psychology sciences. Therefore, the most commonly used measures involve observations, questionnaires, and task performance results. Employing physiological measures has been relatively uncommon in HCI studies mainly due to the high cost of the equipment and the logistical difficulties associated with these measurements. Eye-tracking research has been used extensively both in academia and industry, but other measurements have been rarely used: for example, heart-rate, event-related brain potential (ERP), transcranial Doppler sonography (TCD), position emission tomography (PET), galvanic skin response, and functional magnetic resonance imaging (fMRI). However, in the past 10 years, researchers have been using these technologies to better assess human cognition during HCI tasks (Kramer & Parasuraman, 2007; Lang et al., 2002).

The need for more direct access to human cognition is further strengthened by the inherent biases of indirect measurements. When participants are asked to report their feelings and thoughts, their response is mediated by their self-awareness capability. As a result, subtle changes in their cognitive state may not be

reported, leading to Type II error, which is a failure to find a significant effect when it actually exists. For example, Wilson & Sasse (2000) presented to their participants videos at varying quality levels (manipulated by lower and higher frame rates). They found that, although the participants who were exposed to lower quality video exhibited a significant effect of stress, measured by their galvanic skin response, only 16% of them reported that they were dissatisfied with the quality of the videos.

If a researcher wants to measure the participants' gaze and arousal levels as they interact with online content and advertising, then using tools that record and measure physiological responses is mandatory. The literature reviewed in this section will provide theoretical foundation to the problem domain and a practical understanding of the capabilities and limitations of these measurements.

Eye Tracking

The underlying assumption of eye tracking research is that people process visual information perceived by their eye's retina, also known as the "eye-mind" hypothesis suggested by Just & Carpenter (1976). In their study, they asked participants to mentally rotate a shape and compare the result to one of several possible outcomes. The researchers found a linear relationship between the participants' reaction time and observation time for the two images, which they assumed to be associated with their participants' cognitive processes.

The retina (see Figure 21) is a dense layer of photoreceptors located along the back wall of the eye. These cells collect visual information and send it to the brain to be processed for comprehension (although recent studies show that there

is some degree of crude processing done in the eye level itself (Werbin & Roska, 2007)).

The eye muscles play a crucial part in vision. As one scans visual information,

the eye muscles continuously engage in a rapid, ballistic movement called saccade. During the saccade, humans are not able to perceive new information and are not aware of its existence, a phenomenon named saccadic suppression (Rayner, 1998). Between these movements, there are fixations, in which the eye remains

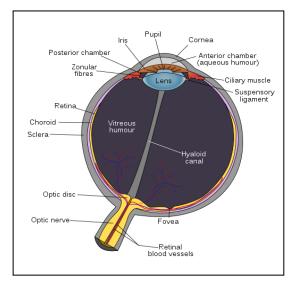


Figure 21. Schematic diagram of the human eye. Source: Wikipedia. Retrieved from http://bit.ly/97RBPt

relatively still on a particular spot for about 200-500 milliseconds. As the eye fixates, most of the light falls in a specific area of the retina called the *fovea*, where about 30,000 color-sensitive photoreceptors are located. This results in the clear image that we perceive in the center of our focal area. Other types of eye movements include *pursuit*, a form of saccade that occurs when the eye follows a target; *vergence*, in which the eyes move inward to fixate on a close target; and *vestibular*, which compensates for head and body movement, allowing a person to maintain their focus on a target as they move.

Early History of Eye Tracking Research

The earliest recorded studies in eye tracking are attributed to Javal (1878; in Wade, Tatler, & Heller, 2003), a French scientist who published a series of articles

studying the physiology of reading. In these publications, Javal noted that humans move their eyes only on a horizontal axis as they read text. In addition, he referred to the rapid accommodation of the eyes as *saccade*. In 1901, American researchers Dodge and Cline developed the first non-invasive eye tracking technique: recording the reflected light from the subjects' corneas as their eyes moved horizontally (Jacob & Karn, 2003). In 1921, Gilliland developed a two-dimensional eye tracker, although only a horizontal or a vertical movement could be recorded at a single time. Further advancement in the technology led to splitting the light reflected from the retina, which enabled the simultaneous recording of the two dimensions (Richardson & Spivey, 2004). Other, more crude techniques, were used by Nixon in 1924 when he examined whether pictorial ads captured the users' attention. In this study, Nixon simply observed his participants' eye movements while reading a newspaper as he hid in a box behind a curtain (Wedel & Pieters, 2007).

So, the bulk of the published eye tracking work in the early 20th century primarily examined the process of reading. It was only in the 1960's and 70's that significant advancements in eye tracking instruments led to a resurgence of eye tracking studies in other problem domains.

The Mechanism of Modern Eye Trackers

Most eye trackers today use the corneal-reflection method (Goldeberg & Wichansky, 2003). These types of trackers consist of an infrared camera mounted beneath a standard computer monitor. The invisible infrared light is directed into the participant's eye, where it is deflected by the retina. In the process, the pupil

appears as a bright oval disc. The deflected infrared light appears on the eye as a small yet distinct glint (see Figure 22) (Poole & Ball, 2000).

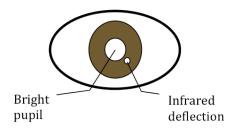


Figure 22. Infrared deflection and bright pupil as seen by the infrared camera.

The eye tracking system is

controlled by software that identifies the center of the pupil and the location of the infrared deflection. The distance between these two reference points is measured and, along with further calibration and trigonometric calculation, the participant's focal point on the screen can be found. While it is possible to calculate the focal point without the pupil location, this requires the use of a chin rest that restricts the participant's head movement. The additional reference point of the pupil location allows the participants to interact freely (with some limitations) with the computer interface (see Figure 23).

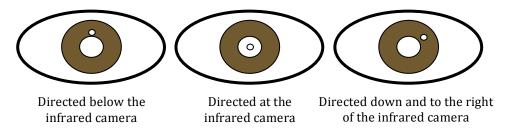


Figure 23. The infrared light deflection glint allows the eye tracking system to locate the participants' focal point.

Eye Tracking Metrics

The most commonly used measurements of eye tracking studies are fixations. The interpretation of fixations is context dependent. In an information retrieval task, a higher frequency of fixations on a particular area, such as an image, may indicate that the target is complex (Just & Carpenter, 1976). In search tasks, a

higher number of fixation clusters may indicate confusion and uncertainty (Jacob & Karn, 2003). Fixation duration is linked to the observer's processing time relative to the focal point (Just & Carpenter, 1976). Longer fixation times are often explained by unfamiliarity with the stimulus (Goldberg & Kotval, 1999).

Another measure, *scanpath*, which consists of saccades and fixations, describes the observer's scanning patterns when exposed to stimuli (Noton & Stark, 1971). For example, a narrow scanpath of an online form would be interpreted as more optimal than a scattered scan pattern (Wroblewski, 2008).

Finally, blink rate and pupil size are used to measure cognitive workload. A high blink rate and a large pupil size may indicate fatigue and high cognitive load (Hyönä, Tommola, & Alaja, 1995). However, these measures are more sensitive to environmental contamination using various ambient light levels, which can introduce additional noise to the data and therefore are less common (Goldeberg & Wichansky, 2003). Table 2 summarizes the common eye tracking measurements used in modern research studies (adapted from Poole & Ball, 2000).

Table 2
List of common eye tracking metrics used in computer interface evaluation.

Eye-Movement Metric	What it Measures
Number of fixations overall	More overall fixations indicate less efficient information retrieval
Fixations per area of interest	More fixations on a particular area indicate that it is more noticeable, or more important, to the viewer than other areas.
Fixations per area of interest and adjusted for text length	If areas of interest are comprised of text only, the mean number of fixations per area of interest should be divided by the mean number of words in the text. This is necessary to separate out: (a) a higher fixation count simply because there are more words to read, from (b) a higher fixation count because an item is actually harder to recognize.
Fixation duration	A longer fixation duration indicates difficulty in extracting information, or it means that the object is more engaging in some way.
Gaze (also referred to as "dwell, fixation cluster" and "fixation cycle")	Gaze is usually the sum of all fixation durations within a prescribed area. It is best used to compare attention distributed among targets. It can also be used as a measure of anticipation in situation awareness if longer gazes fall on an area of interest before a possible event occurs.
Fixation spatial density	Fixations concentrated in a small area indicate focused and efficient information retrieval. Evenly spread fixations reflect widespread and inefficient search.
Repeat fixations (also called "posttarget fixations")	Higher numbers of fixations off-target after the target has been fixated indicate that it lacks meaningfulness or visibility.

Applications in HCI Research

One of the earliest examples of eye tracking data used in HCI research was in 1950, when Paul Fitts investigated the layout of cockpit controls for the US Navy.

Among the tools he used were motion picture cameras to study the movements of pilots' eyes as they used cockpit controls and instruments to land an airplane (Fitts, Jones, & Milton, 1950). However, this work gained little attention at that time.

Instead, Fitts became known for his work in human movement modeling, better known as Fitts' Law (Fitts, 1954).

It was in the 1980's, as personal computers became more prevalent, where when the use of modern eye tracking devices allowed real-time data recording of users interacting with computer displays and freed from physical restrictions. Early studies focused on the development of novel interfaces that allowed disabled users to operate menu commands using eye gazing (Hutchinson, White, Martin, Reichert, & Frey, 1989). Another interesting interface was developed by Tong and Fisher (1984); in their flight simulation system, they created a tracking device that substantially enhanced the resolution around the focal point made by the observer.

Beginning in the 1980's, eye trackers became an indispensible tool to evaluate computer interfaces for the general user (Bolt, 1982; Card, 1984). For example, Card (1984) investigated user interactions with pull-down menus. He found that alphabetically arranged menu commands led to a faster discovery compared with menu commands grouped by function. However, with increased experience, on average, users were able to locate every command with a single eye saccade, regardless of the menu commands' arrangement rule. In addition, Card (1984) observed that dividing menus into distinct groups (boxes), allowed the user to memorize not only the location of a specific command, but also the surrounding commands located in that group.

Cowen, Ball, & Delin (2002) used eye-tracking data to evaluate the usability of commercial cell phones. They asked participants to perform several tasks, such as looking for a particular handset and navigating through various webpages. They

found that a number of fixations and spatial fixation densities indicated issues with website usability.

Eye tracking has also been used to examine aircraft instrument displays (Ottati, Hickox, & Richter, 1999; Thomas & Wickens, 2004). For example, Thomas and Wickens (2004) investigated the scanning patterns of pilots as they interacted with a synthetic vision system. They found that pilots who spent most their time scanning the synthetic display missed unexpected events, such as runway offset. Other pilots who occasionally glanced through windows at the outside world and at other displays had a higher likelihood of coping with these events.

In a classic HCI study, Goldberg, Stimson, Lewenstein, Scott, & Wichansky (2002) evaluated the layout of a new web application. The participants' tasks included renaming a section ("portlet") name, using the employee directory to search for a phone number, and so on. Using the eye tracking measures of fixations, fixation duration, total dwell time, and saccadic amplitude (to define the scanpath), the researchers were able to evaluate the difficulty level of the tasks their participants completed. For example, highly directed scanpaths with few fixations in a search task indicated a low level of difficulty. They also found that participants tended to scan the page section horizontally rather than vertically and scanned the content of the sections before scanning the section header (see Figure 24). The fixation order exhibited by the study's participants led to such design recommendations as placing critical information on the left and the top area of the web page.

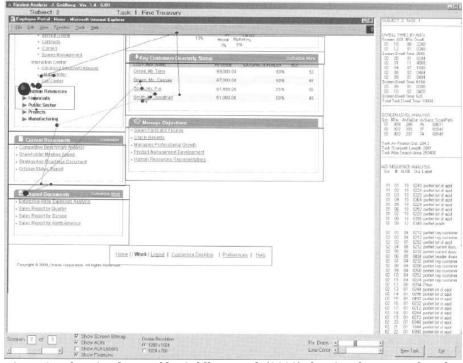


Figure 24. The stimulus used by Goldberg et al. (2002) showing the scanpath and the area of interest

Pan et al. (2004) set an ambitious goal to determine what factors affected eye movement behavior when interacting with online information. The factors they considered were type of task (memorizing a specific piece of information or exploratory browsing), demographic variables of the participants (gender and ethnicity), types of website (business, search, shopping, and news), and order in which the websites were viewed. The eye tracking variables that were considered were mean fixation duration, gazing time, and saccade rate.

Their results showed that users had longer fixation duration on the second page of search and business sites than on shopping and news sites. This was explained by the novelty of subsequent webpages in business and search when compared to the content novelties in shopping and news webpages that introduced contextual cues on their first pages. The gaze time on second pages of search and

news sites was shorter than the gaze time on the second page of business sites. Since a long gazing time was found to be correlated to mental load and complexity (Nakayama, Takahashi, & Shimizu, 2002), the authors explained this result by concluding that business sites demanded more cognitive processing, which increased the participants' mental load. However, when comparing the cognitive load between first to second webpages overall, the authors found conflicting results as data derived from gazing time contradicted the data derived from saccades measurements.

With regard to individual differences, Pan et al. (2004) found that males exhibited longer mean rates of fixation durations than females. Females were found to be more engaged in comprehensive information processing as opposed to males, who focused on fewer areas. An interesting finding by the authors was that participants' scanning behavior of subsequent webpages was somewhat determined by the content and layout of the website's first page; participants implicitly developed viewing strategies for individual websites (regardless of their type) based on the information they encountered on the first page.

Challenges in Eye Tracking Research

As suggested in the findings by Pan et al. (2004), obtaining accurate data from eye trackers can be challenging. Although eye tracking manufacturers have dramatically improved the usability of the equipment, some barriers remain to assured use and accurate measurement. For example, users who wear eyeglasses cannot participate in research studies because the infrared light cannot produce the desired reflection on the cornea. Furthermore, participants cannot communicate

their thoughts during an eye tracking study because verbal communication and facial expression (such as smiling) introduce substantial noise to the eye tracking data. In turn, this makes the task more artificial and may not represent the working environment that the researcher is attempting to simulate.

The amount of data collected by the eye tracker depends on the tracker's sampling rate (normally between 50 to 250 Hz) and the test session duration. These measures commonly yield vast amounts of data that the researcher needs to analyze. Despite of the availability of analysis software such as GazeTracker and Tobii Studio, there are no available standards that define what can be considered as fixation (i.e., can the number of milliseconds that the participant stares at a focal point be considered as a fixation) (Salvucci & Goldberg, 2000).

In addition, dynamic content such as Ajax, pop-ups, and drop-down menus cannot be used with eye trackers because the available software cannot associate the fixations made by the user with the dynamic content.

Finally, interpreting the data can be difficult. As mentioned earlier, conclusions derived from eye tracking analysis are often context dependent. There are a few theories of cognitive load that can lead to overarching interpretations, such as longer fixations on a control element in the interface reflects a participant's difficulty in interpreting the proper use of that control (Jacob & Karn, 2003). However, researchers often must rely on additional quantitative and qualitative data in the form of task performance, post-task think aloud, and observations to interpret the data accurately.

Galvanic Skin Response

Galvanic skin response (GSR), also known as skin conductance response (SCR), is a measure of the electric resistance of the skin. There are two ways to

measure electrodermal resistance: exosomatic and endosomatic. Feré in 1888 first suggested the exosomatic method, in which a small current is passed through the skin. The endosomatic method was introduced by Tarchanoff in 1889; this method measured electrical activity on the skin without the need for externally imposed current. While both methods are used today, the most common method used by researchers is exosomatic (see Figure 25) (Stern, Ray, & Quigley, 2001).

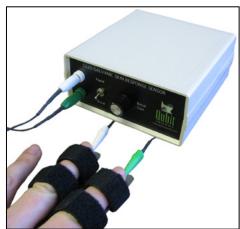


Figure 25. An example of a modern galvanic skin response sensor. Source: Qubit Systems. Retrieved from http://bit.ly/c4x2fK

To measure skin conductivity using the exosomatic technique, a small electric current is passed through two electrodes attached to separate points on the skin. Changes in the muscle and skin tissue affect specific sweat glands called the eccrine glands (located in the palms of the hands and the sole of the feet), which in turn change the GSR values, measured in microsiemens (μ S). These sweat glands are not affected by body temperature and are associated with the participant's psychological state (Boucsein, 1992).

The relationship between electrodermal activity and the central nervous system is unclear. The present standard model was originally proposed by Boucsein (1992), who argued that there are two separate portions of the central nervous

system that affect skin conductivity. The first is called ipsilateral, which includes the hypothalamus, anterior thalamus, and cingular gyrus. The second portion is called contralateral, which contains the lateral frontal cortex (including the premotor cortex), and parts of the basal ganglia. The ipsilateral portion is activated as a response to emotional stimuli while the contralateral portion is activated in response to orienting, cognition, and locomotion.

Early History

The origin of skin conductivity measurements lies in the practice of electrotherapy, which was common in Europe in the mid-1800's (Neumann & Blanton, 1970). In 1879, the French electrotherapist Romain Vigouroux argued that the success of electrotherapy was due in part to changes in the conductivity of the

human body (see Figure 26 for an illustration of the electrotherapeutic technique). In 1888 neurologist Charles Feré linked emotional stimuli to changing levels in the skin conductivity (Neumann & Blanton, 1970; Prideaux, 1920). Feré's work did not attract much attention until it was rediscovered,

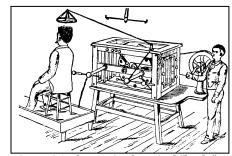


Figure 26. The static electrical "bath." In this version, an electrode over the head directed the discharge, producing static "wind" (Neumann & Blanton, 1970, p. 458).

apparently without prior knowledge, by Swiss electrical engineer E. K. Müller in 1904. Müller communicated his results to his colleague Otto Veraguth, who in turn named the phenomenon "psycho-physical galvanic reflex" (Prideaux, 1920).

In 1906, Carl Jung published his experimental results that showed it was possible to access the human subconscious (in a Freudian sense) by measuring skin

conductivity (Jung, 1906; Stern et al., 2001). In his experiments, he presented his subjects with a list of words and measured the subject's response to individual words. Jung concluded that words associated with high skin conductivity had an emotional impact on the subject's subconscious; these findings were then used to guide the subject's psychotherapy.

For several years there was disagreement among members of the scientific community as to whether changes in skin conductivity reflected changes in emotional response and arousal to a stimuli or were simply random mental activity (Landis & Hunt, 1935). Subsequent studies confirmed a correlation between skin conductivity and emotional response (Theron, 1948; Van Der Merwe & Theron, 1947). Later work established that GSR was effective as a measure of Pavlovian habituation to stimuli (Mundy-Castle & McKiever, 1953). As a result, these studies led to the widespread use of GLS as a measure of arousal—a practice that continues to the modern day.

Applications of GSR in Human-Computer Interaction Research

In a series of studies that examined the effect of slow system response on stress, Kuhmann, Boucsein, Schaefer, & Alexander (1987) found that prolonged response time (8 seconds) resulted in emotional strain that physically manifested as an increase in electrodermal activity. When Wilson & Sasse (2000) evaluated the effect of poor video quality on user satisfaction, they found that showing participants low-quality videos (5 frames/second vs. 25 frames/second) caused increased skin conductivity, indicating stress. However, only four of the 24 participants reported to notice the difference in the video quality. Indicating that the

participants' physical reaction to stimuli did not necessarily accurately reflect their own subjective reporting.

Other studies have attempted to use GSR to depict subtle short-term changes in participants' emotional states. Picard (2000) investigated ways to allow an affective computing system to sense the user's emotional state. She asked her participants to play the computer game DOOM while she monitored their electrodermal activity over the course of several days. The GSR results showed that while the skin conductivity level decreased over time (across days), short-term changes within a game session indicated varying levels of arousal. Lockerd & Mueller (2002) developed a home video capturing system that automatically highlighted salient moments in the scene using physiological sensors. To depict peaks of arousal, they designed a glove with electrodermal sensors connected to the video camera and placed tick marks in the editing file where high arousal moments were detected. Scheirer, Fernandez, Klein, & Picard (2002) took a different approach to detecting participants' emotional response: they deliberately induced highly frustrating scenes during a computer game. This enabled the affective system to differentiate between one emotion and another by using a GSR sensor.

Several published studies employed GSR to evaluate website design and user experience. Ward & Marsden (2003) used a battery of physiological measures, including GSR, to assess the usefulness of such measures as a means to evaluate website usability. They showed participants two screenshots, one of a well-designed interface and one of a poorly designed interface. The participants' task was to find specific information in the website assigned to them and to verbally report what

they found. The study results showed that the ill-designed website induced high levels of arousal compared to the well-designed website. However, the differences between the group means were not significant. In a similar study, Thüring & Mahlke (2007) presented their participants with two different computer-based simulations of a mobile phone. One version was designed to be highly usable while the other was poorly designed. The participants were asked to interact with the interfaces and then answer a usability questionnaire (the Geneva appraisal questionnaire (Scherer, 2001)). Their results showed that participants who were exposed to the poorly designed interface had on average higher levels of electrodermal activity than participants who were exposed to the well-designed interface. These results indicated a higher level of frustration when interacting with an ill-designed interface and the results were correlated with the Geneva questionnaire results.

A few studies have been published where GRS was used for testing in the online marketing domain. Sundar & Wagner (2002) examined the effect of e-commerce website download speed on the users' arousal levels. Using skin conductivity measures, they found that websites with slow download speeds led to higher levels of arousal, which were interpreted as frustration. A carryover effect of the slow download speed was that participants developed a negative attitude to the e-commerce website contents. Sundar & Kalyanaraman (2004) investigated the effect of web advertisement animation on arousal, memory, and impression formation. They found that fast animated ads elicited higher levels of arousal than slow animated ads. However, fast animation did not have a significant effect on ad

recognition (memory) and was generally regarded as less desirable by the users, compared with slow animated ads.

Challenges in GSR Measurement

Electrodermal activity can be easily influenced by room temperature and humidity, which can result in inconsistencies in the data collection. Furthermore, participants' responses may vary as a result of their activities prior to the experiment session and their size, which varies the distance between the GSR electrodes (Ward, Marsden, Cahill, & Johnson, 2002).

As with eye tracking data, there is some difficulty in interpreting the GSR data. Different emotions can produce identical responses. As a result, based on GSR data alone, it is impossible to determine whether a particular response was due to workload, frustration, surprise, or other factors. Therefore, it has been suggested to avoid labeling emotion when using GSR and instead report only the arousal level (Lang, Greenwald, Bradley, & Hamm, 1993).

Cardiac Measures

Heart period (in milliseconds), captured by an electrocardiogram (ECG), is a measure of the time between sequential heartbeats, identified as the interval between successive R wave spikes (see Figure 27). While heart rate

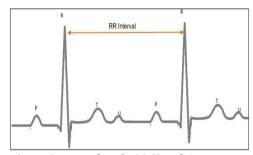


Figure 27. R Peaks of ECG Signal & RR Intervals. Source: http://bit.ly/c0UGA6.

and heart period can be used interchangeably, there is some advantage in using heart period because of a stronger linear relationship between the changes in the

autonomic nervous system and the heart period values, as compared with heart rate value (Berntson, Quigley, & Lozano, 2007).

Contrary to popular belief, a *decrease* in heart rate in fact indicates an *increase* in attention. This reaction is due to activation in the parasympathetic nervous subsystem (one of two branches in the autonomic nervous system) as it is exposed to external stimuli. The stimulation leads to a heightened sense of attention and vigilance, which results in deceleration of the heart rate. The second branch of the autonomic nervous system, the sympathetic subsystem, has an opposite effect: when an external stimulus causes arousal, the sympathetic system accelerates the heart rate. When a stimulus arouses both the sympathetic and parasympathetic subsystems (for example, when exposed to emotional media messages), both acceleration and deceleration signals are sent to the heart. The heart rate is then determined by whichever signal is most dominant. Signals due to cognitive load are usually stronger than arousal and therefore lead to deceleration of the heart rate (Lang et al., 2008).

A third common measure of cardiac activity is heart rate variability (HRV). HRV measures are derived from a spectral analysis of the intervals between consecutive heartbeats (R-R intervals). HRV tends to be suppressed during exposure to stressful stimuli and emerges in a relaxation mode. If a task requires high levels of cognitive resources, HRV is expected to be suppressed (Rowe, Sibert, & Irwin, 1998). The spectral components of the R-R intervals in HRV are divided into three frequencies: very low frequency range, low frequency range, and high frequency range. It was found that the low frequency range (0.04-0.15 Hz) is the

main indicator of mental workload (van Ravenswaaij-Arts, Kollee, Hopman, Stoelinga, & van Geijn, 1993).

Early History

In 1885, the German psychologist Paul Mentz exposed his participants to loud noises and recorded their cardiac reaction (Stevens, 1905). He noticed that his participants' pulse slowed down following the loud stimulus, and that a decrease in the pulse rate was proportionate to the intensity of the sound. Mentz also found that when he asked the participants to solve a series of multiplication problems, their pulse rate increased, which led him to believe that voluntary attention led to an increase in heart rate. A close colleague of Mentz's, Alfred Lehmann, conducted a similar study in 1899 where he noticed an immediate deceleration in the pulse after introducing sudden sensory stimuli. He concluded that this change in cardiac rhythm was associated with an increased demand for attention resources (Darrow, 1929).

However, later studies found opposite effects. Brahn (1901) found that his participants' heart rate increased in response to negative stimuli and decreased in response to positive stimuli. Zoneff & Meumann (1902) found that their participants' heart rate decreased in response to pleasant external and internal stimuli and increased in response to a sudden noise (cited in Darrow, 1929). Darrow (1929) explained these discrepancies by noting that Zoneff and Meumann exposed their participants to a prolonged auditory stimulus as opposed to the quick and sudden auditory onset used by previous studies. In addition, while previous studies

used auditory stimuli, Brahn used tastes and odors as stimuli, which may also have explained the mixed results.

The debate as to which type of stimuli increases or decreases heart rate continued in the literature for several more decades but by the 1960's the accumulated evidence suggested that intense mental activation resulted in a decelerated heart rate and arousal led to an accelerated heart rate. The use of cardiac measures to evaluate emotions (Lacey, Kagan, Lacey, & Moss, 1963) and mental workload (Kalsbeek & Ettema, 1963) proved to be indispensible in psychophysiology research.

Application of Heart Rate Measures in HCI Research

Early studies used cardiac measures to examine mental workload in human factors research (see reviews by Kramer, 1991; Kramer & Raja Parasuraman, 2007). Sirevaag et al. (1993) carried out one of the earliest studies that utilized HRV to examine mental workload while interacting with digital displays. Their study used a flight simulator to examine the difference between digital and verbal displays. They asked a group of trained US Army helicopter pilots to simulate a reconnaissance mission in enemy territory. For each mission, the pilots were instructed to identify and report the location of enemy positions and their time of arrival to these locations. The pilots were prompted to report their findings to base via either digital keypad or verbal update. In addition, the workload was manipulated by changing the number of enemy positions and other obtrusions. Their results showed that using verbal communication in a high load condition resulted in higher HRV values than when using a digital display. However, the HRV results were not conclusive, as

the results could have been contaminated by lack of adequate control over the amount of data that was voluntarily transmitted by the pilots in each of the conditions.

Tattersall & Hockey (1995) also examined workload in a flight simulator using both HRV and heart rate (HR) measures. The pilots' task was to perform a routine flying mission. During the flight the researchers introduced several bugs into the simulation software; the pilots were responsible for detecting the bugs, performing diagnostics, and fixing the bugs. The bugs were introduced in three flight stages: take-off, level flight (cruising), and landing. The mental load was manipulated using a varying number of bugs for different flight stages. Their results showed that HRV values while performing maintenance activity during *non*-routine flight were significantly lower than routine flight activities. In addition, while HR was elevated during take-off and landing stages, they did not find a significant effect for variation in mental load.

Rowe et al. (1998) examined HRV values as their participants engaged in an air traffic control game. The researchers manipulated the participants' mental workload using varying numbers of roaming airplanes on the display (higher numbers of airplanes induced a greater mental workload). Five out of the 13 participants had previous experience with air traffic control. The study results showed a marginal effect for the number of planes observed across the entire group. But analysis of the experienced participants revealed a significant sensitivity: as the number of planes on the monitor increased, the experienced participants' HRV values decreased; but during a high workload level of 12-16 planes, their HRV

increased. The authors explained this result by arguing that in a very high level of workload the task became too difficult, which led the participants to disengage from the task.

In addition to assessing mental workload, cardiac measures have been used to examine orienting response as a measure of attention. Lang et al. (2002) used heart rate data to evaluate orienting response to text (news headline), boxed text (news headline in a bordered box), warnings (computer prompts/pop-ups), banner ads, and animated banner ads. The authors found that only warnings and animated banner ads elicited a significant orienting response; these were also the only stimuli recognized by the participants in a subsequent memory test.

Diao & Sundar (2004) used heart rate to measure the orienting response to pop-ups and banner ads while controlling for animation in both of these stimuli (see Figure 28). During the experiment, the participants were asked to read, at a normal reading speed, the content of the web pages presented to. The participants were not able to interact with the page (i.e., scrolling or clicking back and forth); screenshots of the pages automatically rotated every few seconds. The authors found that pop-ups led to a greater orienting response than banners (BPMs for pop-ups were lower).



Figure 28. An example for the stimuli used in Diao and Sundar (2004).

With results similar to Lang et al. (2002), ad recall rate was greater for popups as compared to banners. But the authors failed to find a main effect for animation. They explained this result by hypothesizing that the combination of banners and pop-ups on a single webpage may have led to negation of the task (similar to Rowe et al. (1998)). In addition, because animation was present on websites commonly used by the college age participants, habituation may have influenced on these results.

Finally, an interesting study by Wise & Reeves (2007) examined the effect of user control over orienting response. The authors presented the participants a series of both emotional and neutral images and asked them to rate their arousal level using a standard questionnaire. The authors manipulated user control of the image scrolling. On half of the trials, the computer automatically rotated the images; on the other half of the trials, the participants had the freedom to switch between the images at will. The study results showed that orienting response, as measured

by heart rate, was greater for emotional images that were controlled by the *computer*. In addition, the participants rated the images that were controlled by the computer as more emotionally provoking than images that they could control.

Limitations of Cardiac Measures

There is no standard method to measure psychophysiological effects using cardiac techniques. As observed in the literature reviewed in this section, some authors chose HRV while other chose heart rate. Oftentimes, the rationale behind choosing one method or another is absent or ambiguous.

As with other physiological instruments, cardiac measures require the participants to be hooked-up with obtrusive electrodes that may increase their discomfort and anxiety. Furthermore, intake of caffeine and individual biases may also contaminate the data and result in missed or skewed effects.

The breadth of the literature reviewed in this section confirms the benefits of using physiological measures in HCI research. However, the logistical challenges of these measures (mainly cumbersome and noisy data collection) have led many researchers (based on personal communication) to conclude that these measures are neither efficient nor cost-effective. Yet, as more studies are published, researchers continue to find effective coping techniques to mitigate these challenges. For example, to reduce data noise, Mandryk, Inkpen, & Calvert (2006) proposed to normalize individual data prior to the statistical analysis. In addition, it is a common practice to start recording the data after allowing the participants 5-10 minutes to interact with the system and feel more comfortable.

Finally, most (if not all) the studies cited here used a *battery* of measures to test their hypotheses. This battery often consisted of a combination of questionnaires and multiple physiological measures. There are two main reasons why this approach is advisable. First, the diversification of tests ensures that a failure in one measure will not risk outcomes for the entire project. Second, and more importantly, diversification allows the researcher to correlate the results of one measure to another, thus ensuring the validity of their interpretation to the findings.

Summary

For this dissertation, I have investigated the problem domain of ad avoidance. While ad avoidance could be affected by a large set of behavioral and design factors, I focused my research on three topics that have received little or no attention in the academic literature:

1. The effect of search type. Previous studies found that known-item search is qualitatively different from exploratory search. While users who engage in a known-item search are scanning the text for a particular item, users who engage in an exploratory search are often integrating multiple sources of information to answer a query. Due to variability in experimental protocols, previous studies have not reached a consensus that explains the differences in ad perception as a function of search type. Based on load theory, a high perceptual load would predict that users engaged in a known-item search are less likely to notice advertisements as compared to users

engaged in an exploratory search. This is because the latter users are affected by cognitive load, which leaves them more vulnerable to distractors.

2. The effect of saliency. Highly salient ads have not improved the performance of online banner ads, in contrast to a growing trend of rich media usage in online ads. CTR continues to plummet, no matter whether the ad features animation, bold colors, or pop-ups.

Based on a few studies that examined the role of distractors in visual search, salient distractors may in fact be less effective as evidenced by the lower search times in tasks where salient distractors were present. Therefore, highly salient ads may promote ad avoidance as they serve as visual guides for where users should *not* be looking.

However, the factor of saliency is likely to play a different role in known-item vs. exploratory search; users who engage in a known-item search are more likely to notice low-salient ads, while users who engage in an exploratory search are more likely to notice highly salient ads.

3. The effect of repetition. Due to conflicting findings reported in the literature, it is currently unclear what effect repetition has on banner ad perception. There are two conflicting models: the asymptotic habituation model and the habituation-tedium model.

The asymptotic habituation model predicts a decrease in orienting

response to banners as a function of repetition. Therefore, with each subsequent exposure to a banner ad, the user pays less attention to the ad.

The tedium-habituation model predicts that early exposure to a banner ad is likely to attract the users' attention due to the users' uncertainty with regards to the nature of the new object. However, with increased exposure, habituation plays a positive effect by decreasing this tension, which leads to an increased orienting response to that banner ad. Finally, as the number of repetitions increases, tedium takes effect, which leads to a decrease in the attention given to subsequent banner ads.

Although both models are supported by previous research, it is possible that while repetition decreases the orienting response to ads, it may support long-term retention of the brand or product name.

In addition to these factors, I examined the use of physiological measures in HCI research—specifically, eye tracking, galvanic skin response (GSR), and cardiac measures. With the exception of eye tracking, GSR and cardiac measures are rarely used to measure users' engagement with online advertising and are seldom used in HCI research. This is in spite of evidence that suggests that physiological measures could detect minute effect sizes, which would otherwise remain undiscovered. Therefore, assessing the effectiveness of these measures and their correlation to one

another can advance not only the study of ad avoidance, but benefit the HCI research community in general.

Hypotheses

Based on the research question and preceding literature review, the following hypotheses were tested:

- 1. There will be a main effect for search type, such that orienting response to banner ads for participants who engage in a known-item search task will be lower compared to participants who engage in an exploratory search task.
- 2. There will be a main effect for ad saliency, such that the orienting response for low salient ads will be greater than medium and highly salient ads.
- 3. There will be an interaction between saliency level and search type on the orienting response to banner ads. The orienting response to highly salient banner ads will be significantly **lower** than low salient ads among participants who engage in a **known-item** search. In contrast, the orienting response to highly salient banner ads will be **greater** than low salient ads among participants who engage in an **exploratory** search.

In addition to the preceding hypotheses, I will attempt to resolve the tension between the asymptotic habituation model (AHM) and the habituation-tedium

model (HTM). Since neither of these models have been evaluated in the context of online advertising, no specific hypothesis can be made.

STUDY 1—EXPERT USERS

Experimental Design

To test these hypotheses, I used a 2x3 mixed factorial design using two independent variables: Search Type (known-item vs. exploratory), which was manipulated within subjects, and Ad Saliency (high, medium, or low), which was manipulated between subject groups.

The primary dependent variable is the orienting response to banner ads. This variable was measured using the number of eye fixations, covert attention (using the Stroop test), and direct attention (using of recall and recognitions questionnaires). To examine whether our manipulation of the banner ads had any effect on the user experience of the website, I included an "attitude towards ads" questionnaire (MacKenzie, Lutz, & Belch, 1986).

Participants

Based upon a preliminary power analysis, 51 participants (34 females and 17 males) were recruited. All participants were students attending the University of North Carolina at Chapel Hill (mean age =22.2, SD=4.44). Three participants were dropped from the final analysis due to severe technical issues during data collection.

The inclusion criteria were:

- 1. Able to read and communicate in English.
- 2. 18 years old and above.
- 3. Unaided vision (did not wear glasses or contact lenses)⁵.

In exchange for participation in this study, each participant received \$15.

Participants who could not complete the study due to technical malfunctions received \$10.

Materials

Hardware

The participants used a computer workstation running Microsoft Windows XP with a display resolution of 1280x800 pixels. The eye gaze data was collected using an ASL D6 60Hz eye tracking system and recorded using GazeTracker v.9.0.7000.1000. The experimenter sat behind the participant and used a second workstation to monitor and control the eye tracking system (see Figure 29).

_

⁵ As mentioned in the literature review, modern eye trackers rely on the reflection of infrared light to identify the gaze direction. Eyeglasses or contacts break the infrared beam, which significantly inhibits the eye tracker's operation.



Figure 29. Hardware setup in the interaction design lab.

Stimuli

Website

Google Sites was used to create a website providing travel information for 10 countries around the world. The home page included instructions for the participants and images of backpackers to help convey the notion that the participants were planning to travel abroad (see Figure 30).

Instructions

The purpose of this section is to examine how potential visitors retrieve information about their destination using articles as their primary source.



Imagine that you were given a free plane ticket to one of 10 destinations around the world. Based on your search results, you found 5 articles that describe major tourist attractions in every country.

For each country, you will be asked to answer different questions based on the articles presented in the results pages of that country. In some cases you will be asked to rate the **perceived level of safety** for traveling in that country while in others you will be asked to answer **specific questions about the tourist attractions** that were presented in the results page.

Please do **NOT** use the Ctrl+F function to find specific keywords. Also, do **NOT** navigate to individual search result by opening it on a new tab, you must navigate back an forth using a single browser window. If you have any questions about the experiment please ask them **NOW**.



Figure 30. Instruction/home page of the website.

The information for each country included five search results. Each search result was linked to an article about a popular tourist destination in that country. As part of their tasks, the participants were required to answer a series of questions: either five known-item search questions⁶ (one for each search result) or one exploratory question. See Appendix A for the complete search tasks.

For example, for Bhutan, one of the known-item search questions was:

Which king does the National Memorial Chorten honor?

- A. Tenzin Drukgye
- **B.** Jigme Dorji Wangchuck
- C. Mipham Wangpo
- **D.** Chhogyel Sonam Gyaltshen

The exploratory question for Bhutan was:

The Bhutanese measure their quality of life using "gross domestic happiness." Based on the articles you read, please postulate why they chose this measure as

⁶ To ensure that the known-item questions will be challenging to answer, 7 PhD students were asked to answer the questions without gaining access to the articles. The mean score in percentage of correct response was 26.57, SD=4.43. Two questions were replaced due to a high ratio of correct responses (71.43% and 85.71%).

opposed to the standard income measure. Does this measure describe their quality of life more accurately?

The sequence of search tasks (known-item vs. exploratory) was predefined using counterbalanced distribution across the 10 trials (one trial per country). To control for potential order bias of the search type, two identical versions of the website were created with different sequences of search tasks (see Table 3). In addition, to control for an order bias of the countries, 10 country sequences were created using a Latin Square.

Sequences of search tasks (KI=known-item; E=exploratory)

Table 3

	South Korea	Bulgaria	Belarus	Moldova	Paraguay	Guatemala	Mauritania	Turkey	Djibouti	Bhutan
Sequence 1	KI	Е	KI	Е	KI	Е	E	KI	E	KI
Sequence 2	E	KI	E	KI	Е	KI	KI	E	KI	E

To minimize disruption to the eye tracking data collection, the results page included those questions, where participants could submit their answers electronically (see Figure 31 and Figure 32).

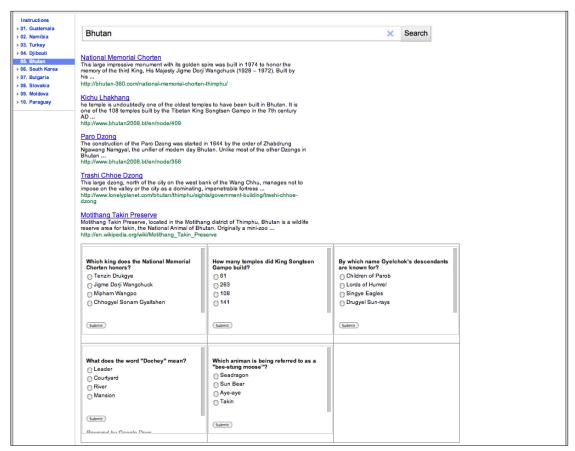


Figure 31. "Search results" page for Bulgaria with known-item search tasks.

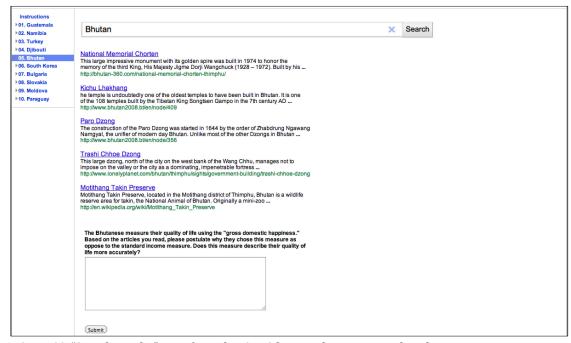


Figure 32. "Search results" page for Bulgaria with an exploratory search task.

As mentioned earlier, each country's main page had five sub-pages corresponding to the search results. The layout of the webpage was inspired by a typical page from The New York Times website (see Figure 33):

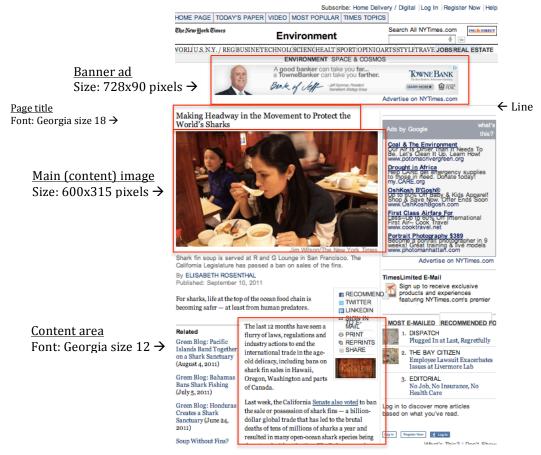


Figure 33. An example of a New York Times article layout.

The content of each sub-page was divided into three paragraphs to control for answer location in the known-item search condition (top, middle, or bottom). The answer location was counterbalanced across all the articles. Each sub-page also included a content image embedded in the middle paragraph on the left side of the page. That image corresponded to the tourist attraction being reviewed. Each sub-page included a banner ad above the first paragraph. Finally, the length of the articles was adjusted to about 300 words and the entire page was resized to 75% so

that the participants would not need to scroll down the page. This layout ensured that the banner ad was always visible (see Figure 34).



Figure 34. Example of sub-page design.

Banner Ads

The brands for the banner ads were all fictitious to avoid a potential confound due to brand recognition. The levels of saliency were defined as follows:

- 1. High = animated banner ad where both the image and the tag line moved in a loop as long as the banner ad was visible (see an example at https://sites.google.com/site/study1hsseq6a/bhutan/national-memorial-chorten). The animation format was identical across all the banner ads used in this study.
- 2. Medium = Static (non-animated) color banner ad.
- 3. Low = Static grey-scale banner ad with a white background and black font.

Five brands were created: VIP Car Rental, Omni Airlines, Marion Hotels, Sugarland Travel Agency, and Tempah Resorts. The chosen banner dimensions

followed the standard "leaderboard" format of 728x90 pixels, which is the second most effective banner ad format used in the industry⁷ (Go, 2008; Kaiser, 2007). However, since the webpage was resized to 75%, the actual dimensions of the banners were 546x68 pixels. Each banner consisted of five elements: background color, brand name, slogan, image, and a call for action ("click here") (see Figure 35). The banners were distributed equally across each country's sub-pages in a random order within each country's search results.



Figure 35. The five banner ads used in this experiment.

Procedure

The experiment took place at SILS' Interaction Design Lab (iDL). Upon arrival, the participants were seated approximately 50 centimeters from the workstation display. Each participant was asked to read the information sheet

⁷ The most effective banner ad is a rectangular (300x250 pixels) embedded ad, which is qualitatively different from common banner ads. Embedded ads are effective because they replace an area where text would normally be found with an ad, thus forcing the user to process the image.

(Appendix B) and sign the consent form if they wished to continue the experiment (Appendix C). The experimenter repeated some of the information that was presented in the information sheet and briefly reviewed the navigation structure of the website. Before continuing to the next phase, the experimenter encouraged the participants to ask questions about the study. After answering the participants' questions, the experimenter began calibrating the eye tracker. Following a successful calibration, the participants were instructed to begin their tasks. After completing five trials (5 countries out of 10), the participants took a five-minute break and then completed the remaining trials (see Figure 36).

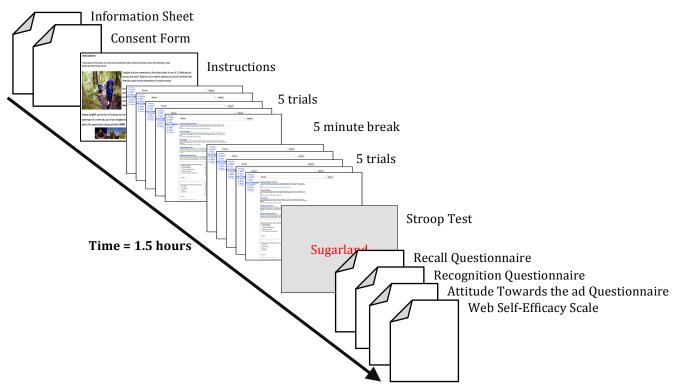


Figure 36. Illustration of the experimental procedure.

Post-test Questionnaires

After taking part in the experiment, the participants were asked to complete several tasks ad questionnaires to examine whether they retained any information about the banner ads and about their attitude towards the ads.

Covert Attention

Using the Stroop test (MacLeod, 1992; Stroop, 1935), I examined whether the participant covertly noticed the banner ads without being consciously aware that they were exposed to the banner ad content. The software application for the Stroop test was developed specifically for this study. Using three keyboard keys (left arrow, down arrow, and right arrow), the participants were asked to indicate as fast as they could the font color (red, green, or blue) of the word that was presented to them. When the correct key was pressed, the next word appeared. If a wrong key was pressed, the word did not change until the correct key was hit.

The participants were exposed to two sets of words that were repeated five times in a random order and in a random font color. Each set included 10 words. One set, "target", contained words that appeared in the banner ads: VIP, Omni, Airlines, Tempah, Marion, Sugarland, Hotel, Airline, Car, Resort, and Travel. Another set, "non-target", included words that did not appear in the banner ads. Instead, those words appeared later in the direct memory questionnaires to prevent a potential carryover effect. These words were Station, Eldan, Abigail, El San, Maupin, Engiadina, House, Walk, Bicycle, and Train. The reaction time for each word was recorded with the assumption that if the participants had covertly processed the banner ads, then the reaction time for the target words would be *longer* due to additional mental processing (MacLeod, 1992).

Direct Attention

After the Stroop test, two memory questionnaires were administered based on Kalyanaraman (in press). Both questionnaires included 15 questions about the banner ads (three questions for each banner ad). The first questionnaire included recall questions, such as "Name the company that appeared in the [travel agency] ad:_______" (see Appendix D). The second questionnaire included the same questions as the first, except that they were multiple-choice questions. For example, "Name the company that appeared in the [travel agency] ad: Kalos, Maupin, Sugarland, or Travelink" (see Appendix E). Each correct answer was given one point, so the scale for each questionnaire was 0-15.

User Experience Questionnaire

Participants' attitudes towards the banner ads were measured using a standrdized questionnaire developed by MacKenzie, Lutz, and Belch (1986)(see appendix F). Participants were asked to rate 14 items on a 7-point Likert scale. These included items such as appeal, informativeness, excitement, interest, and others.

Web Self-Efficacy Scale

Previous studies suggested that web experience may affect ad avoidance behavior (Cho & Cheon, 2004; Danaher, Mullarkey, & Essegaier, 2006). Therefore, to control for potential confounds, I administered the web self-efficacy questionnaire (WUSE) (Eachus & Cassidy, 2006; Eachus, Hogg, & Cassidy, 2006). Earlier development of the computer user self-efficacy (CUSE) tool by Cassidy & Eachus (2002) showed computer user self-efficacy is highly predictive of computer

competency and experience. Based on these findings, Eachus and Cassidy (2006) extended CUSE to the World Wide Web. WUSE is one of the few validated scales that examines self-efficacy beliefs among adults. Furthermore, the reported reliability of WUSE is exceptionally high (Cronbach's Alpha⁸ = 0.96). However, I modified some of the questions so they better reflected modern web behaviors and use. For example, instead of asking participants about their experience with the use of FTP software to upload and share images, they were asked about their experience with the use of social networks to share photos⁹. Following the completion of the data collection, an analysis of reliability was conducted on the modified version of WUSE scale, which showed a Cronbach's Alpha of 0.72 (N=48). See Appendix G for the items that appear in WUSE and an explanation about the analysis.

Results

Gaze Data

The eye tracking data was collected using GazeTracker. For this analysis, a fixation was defined as collection of discrete gaze points within 40 pixels diameter during *at least* 200 milliseconds. Each webpage included a "LookZone" that defined the location of the banner ad. The dimensions of the LookZone were 617x123 pixels

_

⁸ Cronbach's Alpha is a measure of internal reliability in psychometric tests. It ranges from zero to 1, with a minimum acceptable value of 0.7 (Field, 2009).

 $^{^9}$ Prior to deploying the study, I asked 10 Ph.D. students at SILS to complete the new version of WUSE, in addition to reporting the number of hours they use the web daily. The average score was 85.24 and there was a high correlation between the WUSE results and their self-reported number of hours of web usage: r= .6, p<0.05.

(see Figure 37)¹⁰. The location of the LookZone and its dimensions remained constant across all the webpages and the experimental conditions.



Figure 37. Example of a LookZone used throughout the data analysis.

In cases where the calibration was not accurate, the gaze data had to be manually adjusted. When the gaze data was adjusted, the webpage structure was used as an orienting guide for the true location of the participants' fixations. In addition, the degree of change was recorded in the analysis log, and the change was applied across all the webpages within that recording session (see Figure 38).

¹⁰ The ASL D6 eye tracker has an accuracy of 0.5 degrees. Therefore, the dimensions of the LookZone included fixations within 30 pixels of the banner ad.

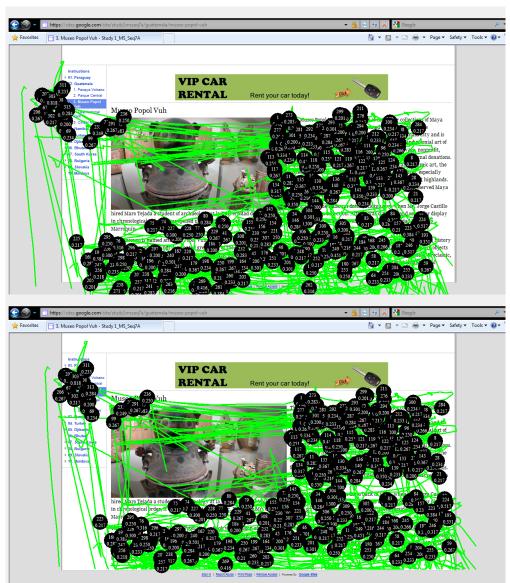


Figure 38. An example of a manual gaze shift. The top image shows the original gaze data and the bottom image shows the adjusted gaze data. In this case, the gaze data was shifted 60 pixels to the right and 20 pixels upwards.

The gaze data was exported to Microsoft Excel spreadsheets, in which the results for each country were averaged across five webpages (each country contained five web articles). Then the data was imported into an SPSS sheet along with the post-test questionnaire results.

Statistical Assumptions

Prior to conducting the analysis of variance, the data obtained through the experimental manipulation needed to comply with three assumptions (Maxwell & Delaney, 2004):

- 1. Normal distribution: tested using the Kolmogorov-Smirnov test.
- 2. Homogeneity of variance across all the experimental groups: assessed using Levene's test.
- 3. Observed scores must be statistically independent of each other: controlled through the randomization of the independent variables and the use of multiple sequences of search types. A t-test between the results obtained following sequence 1 to sequence 2 should not be significant.

The first analysis of the "number of fixations" data was found to violate the normal distribution assumption in three out of six experimental conditions (see Table 4).

Table 4

Test of normal distribution

	Condition	Kolmogorov-Sn	Shapiro-Wilk		
		Statistic	df	Statistic	df
Known_Item_Fixation	Low Salience	0.171	16	0.944	16
	Medium				
	Salience	0.239*	16	0.764**	16
	High Salience	0.211*	16	0.783**	16
Exploratory_Fixation	Low Salience Medium	0.184	16	0.891	16
	Salience	0.163	16	0.916	16
	High Salience	0.237*	16	0.887*	16

Note. *p<0.05, **p<0.01

Although the box plot showed a few cases that could be outliers (see Figure 39), further analysis did not find any evidence to confirm that supposition.

Therefore, the entire data set was transformed used the square root method ($\sqrt{X_i}$) (Field, 2009, p. 80)¹¹.

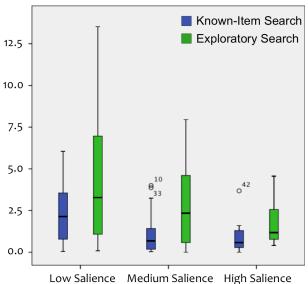


Figure 39. Box plots of the number of fixations data.

Following the data conversion, another test of normality was conducted. This test found that the distribution of the converted data was *not* significantly different from a normal distribution (see Figure 40).

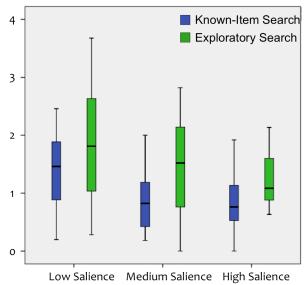


Figure 40. Box plots of the transformed number of fixations data showing normal distribution.

¹¹ The significance level of the ANOVA did *not* change as a result of this data conversion.

To test the second assumption, homogeneity of variance, a Levene's test was performed on the transformed data. The results showed that the variance between the salience levels in known-item and exploratory search conditions was not significantly different: $F_{(2,45)} = 0.13$, ns; $F_{(2,45)} = 2.72$, ns, respectively.

Finally, an independent t-test analysis between the two sequences showed no significance difference for both the known-item search and the exploratory search conditions: $t_{(46)}=0.29, ns;\ t_{(46)}=0.39, ns$, respectively.

Exclusion of Potential Confounds

In addition to validating the statistical assumptions, it was important to exclude potential confounds. I examined specifically the effect of age, web self-efficacy, sex, and answer location on number of fixations across search type and ad saliency levels.

Age and Web Self-Efficacy

Based on a linear regression analysis, there was no significant effect of age and web self-efficacy scores on average number of fixations: $F_{(2,45)}=0.11$, ns., r=0.022.

Sex

Based on the Kruskal-Wallis test, there was no significant difference between the average number of fixations of males (M=2.98, SD=2.60; n=16) and females (M=1.81, SD=1.58; n=32): $\chi^2_{(1)}=1.87, ns., \eta^2=.043$.

Answer Location

The analysis of answer location (1st, 2nd, and 3rd paragraph) would have required a time-consuming manual scan of the entire data set (2400 data points). To

speed the analysis, a one-way repeated-measured ANOVA was performed over a subset of 9 participants (first 3 participants in each ad salience group → 450 data points) followed by a power analysis to determine whether additional analysis is needed. Based on the power analysis, it would have taken 150 participants (three times the number of participants recruited for this study) to find a significant main effect for answer location (assuming such an effect existed). Therefore, it was concluded that the likelihood of the answer location affecting the study results was minimal and should not be taken into account in the statistical analysis.

Analysis

Number of Fixations¹²

To examine the effect of ad salience and search type on number of fixations, a mixed factorial 2x3 ANOVA was conducted.

There was a significant main effect for search type on the number of fixations on the banner ads: $F_{(1,45)}=39.31, p\leq 0.01, partial \eta^2=.466$. When conducting known-item searches, the participants looked *less* at the banner ads (M=1.44,SD=1.48) when compared with exploratory searches (M=2.96,SD=2.88).

In addition, there was a significant main effect for ad salience on the number of fixations on banner ads: $F_{(2,45)}=3.38, p\leq 0.05, partial\ \eta^2=.131$. The Tukey HSD *post hoc* test revealed that when exposed to highly salient ads (color+animation), participants looked *less* at those ads (M=1.32, SD=1.08) when

 12 All the statistical tests for number of fixations were conducted using the transformed data. However, the means and standard deviations represent the original values.

93

compared with low salient ads (non-animated black & white) (M=3.26, SD=2.80) ($p \le 0.05$). No significant difference was found between the medium ad salience level (non-animated color) to other levels of ad saliency (M=2.04, SD=1.93).

Finally, there was no significant interaction effect between the search type and the ad salience level: $F_{(2,45)}=0.33, ns.$, $partial\ \eta^2=.014$ (see Figure 41).

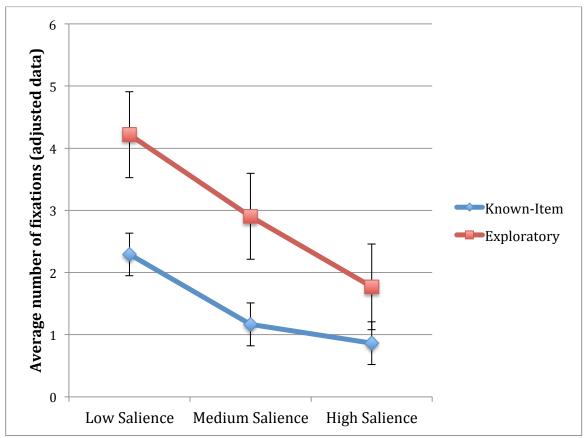


Figure 41. The effect of ad saliency and search type on number of fixations. Error bars represent standard error.

Number of Fixations as a Function of Repetition

Prior to the analysis, the fixation data in each trial had to be transformed to a normal distribution using the square root method. In addition, to decrease the variability in the data, the trials were consolidated into five samples by averaging pairs of sequential trials ((1+2)/2, (3+4)/2...(9+10)/2). Using repeated measures

ANOVA, there was no significant effect of trial order¹³ on number of fixations: $F_{(4,188)}=1.83$, ns., $partial\ \eta^2=.038$. In addition, using a Curve Estimation

procedure, there was no significant model fit; see Table 5 and Figure 42.

Table 5

Model fit results for the effect of repetition on number of fixations

 , ,,	, .				
Curve	R Square	F	df1	df2	Sig.
Inverse	.000	0.01	1	238	0.91
Ouadratic	.007	0.79	2	237	0.46

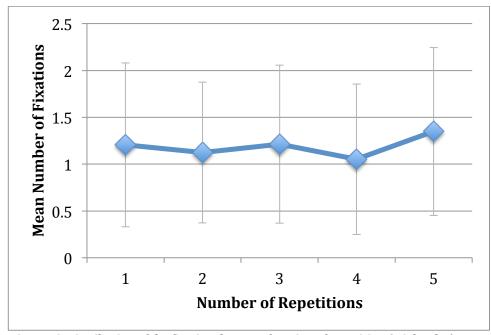


Figure 42. Distribution of the fixation data as a function of repetition (trial order).

Other Gaze Data Analysis

In addition to number of fixations, other gaze variables were analyzed. These were:

1. Total time that the banner ad was gazed at ("Time")

¹³ One trial represents an average number of fixations across a single country.

- 2. Number of repeated gaze fixations on the banner ad ("Gaze Repetition")
- 3. Percent of fixations on the banner ad ("Percent of Fixations").

All of the data for those variables had to be transformed and analyzed using the same method used for number of fixations.

As expected, the results for those variables were very similar to those obtained for number of fixations. Nevertheless, since they added another qualitative dimension to the interpretation of the fixations results, I include them here.

Time

There was a significant main effect for search type on the time that the banner ad was observed: $F_{(1,45)}=35.04, p\leq 0.01, partial \eta^2=.438$. When conducting known-item searches, the participants spent *less* time looking at the banner ads (M=0.90~sec., SD=1.05) as compared with exploratory searches (M=1.83~sec., SD=1.69).

In addition, there was a significant main effect for ad salience on time looking at the banner ads: $F_{(2,45)}=3.91, p\leq 0.05, partial\ \eta^2=.148$. The Tukey HSD *post hoc* test revealed that when exposed to low salient ads, participants spent *more* time looking at those banner ads (M=2.09,SD=1.73) as compared with medium salient ads (M=1.03,SD=0.99) and high salient ads (M=0.98,SD=0.97) $(p\leq 0.05)$. No significant difference was found between the medium and high ad salient ads.

Finally, there was no significant interaction effect between search type and ad salience level: $F_{(2,45)}=0.11, ns.$, partial $\eta^2=.005$.

Gaze Repetition

There was a significant main effect for search type on the number of times that the participants shifted their gaze from the content area of the webpage to the banner ad: $F_{(1,45)} = 42.44$, $p \le 0.01$, $partial \eta^2 = .485$. When conducting knownitem searches, the participants shifted their gaze towards the banner ad *fewer* times (M = 5.54, SD = 4.1) as compared with exploratory searches (M = 9.27, SD = 6.48).

However, there did not appear to be a significant main effect for the impact of ad salience nor an interaction effect between ad saliency and search type: $F_{(2,45)}=2.8, ns., partial \ \eta^2=.111; \ F_{(2,45)}=0.22, ns., partial \ \eta^2=.010,$ respectively.

Percent of Fixations

There was a significant main effect for search type on percent of fixations on banner ads, $F_{(1,45)} = 9.32$, $p \le 0.01$, $partial\ \eta^2 = .172$. When conducting knownitem searches, the percentage of fixations on banner ads relative to the entire webpage was $smaller\ (M=4.33\%, SD=4.96)$ as compared with exploratory search (M=6.00%, SD=7.23).

In addition, there was a significant main effect for ad salience on percent of fixations: $F_{(2,45)} = 4.68$, $p \le 0.05$, $partial \eta^2 = .172$. The Tukey HSD post hoc test revealed that when exposed to low salient ads, the percentage of fixations on banner ads was greater (M = 8.63%, SD = 8.59) as compared with medium salient ads

(M=3.48%, SD=3.60) and high salient ads $(M=3.4, SD=3.12)(p \le 0.05)$. No significant difference was found between the medium and high ad salient ads.

Finally, there was no significant interaction effect between the search type and the ad salience level: $F_{(2,45)}=0.85$, ns., $partial\ \eta^2=.004$.

Post-test Questionnaires

Statistical Assumptions

The distribution of the Stroop test, recall, and recognition results was not normal; the only exception were the results for the participants' attitude to the banner ads, which did follow a normal distribution However, while it was possible to transform the Stroop test and recognition scores to a normal distribution (using square root transformation), the data for the recall questionnaire could not be transformed due to a floor effect (see Figure 43). Therefore, to analyze the effect of ad saliency on recall, I used a non-parametric test.

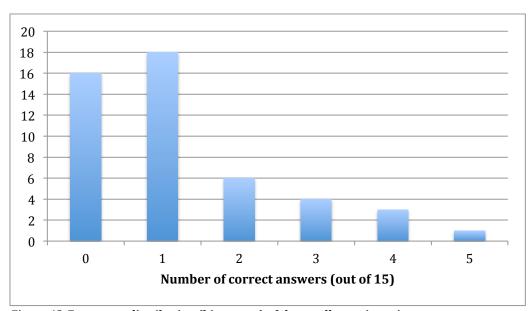


Figure 43. Frequency distribution (histogram) of the recall questionnaire scores.

Stroop Test

Based on a paired t-test analysis, there was no significant difference between the target and non-target words in the low salience condition, $t_{(15)}=-0.77, ns.$, r=.2; medium salience condition, $t_{(15)}=0.54, ns.$, r=.12; and high salience condition, $t_{(15)}=-0.62, ns.$, r=.16; see Table 6.

Table 6

Descriptive statistics for the Stroop test analysis (all units are in milliseconds).

	Low Salience		Medium Salience		High Salience	
	М	SD	М	SD	М	SD
Target Words	704.5	97.76	692.88	82.39	730.38	92.87
Non-Target Words	711.82	98.11	687.44	111.85	740.25	97.27

Recall

Based on the Kruskal-Wallis test, there was no significant difference in average number of correct answers in the recall questionnaire between low salient ads (M=1.00,SD=1.03), medium salient ads (M=1.5,SD=1.67), and high salient ads (M=1.19,SD=1.11): $\chi^2_{(2)}=0.427,ns.$, $\eta^2=.009$. In addition, there was no significant correlation between recall scores and the average number of fixations (nor with any other gaze measure): r=.121,ns.

Recognition

Although the mean recognition scores for medium (M=5.94, SD=2.95) and high salient ads (M=5.75, SD=2.43) were higher than the low salient ads (M=5.06, SD=2.72), that difference was not significant: $F_{(2,45)}=0.4, ns.$, $partial\ \eta^2=.017$. In addition, there was no significant correlation between recognition scores and the average number of fixations (nor with any other gaze measure): r=.202, ns.

Attitude Toward the Banner Ad

Based on one-way ANOVA, there was no significant difference in the average rating of the attitude towards the ad questionnaire between low salient ads (M=2.77,SD=0.71), medium salient ads (M=3.04,SD=0.99), and high salient ads (M=3.22,SD=0.76): $F_{(2.45)}=1.23,ns.$, $partial \ \eta^2=.052$

Summary

The results obtained from the gaze data supported the first two hypotheses:

- Exploratory searches led to a higher orienting response toward banner ads as compared to known-item searches.
- Low salient ads led to a higher orienting response as compared to highly salient ads.

However, I did not find evidence for an interaction effect between ad salience and search type. In addition, I did not find a trend in the gaze data as a function of repetition. Therefore, there was no evidence to support either the asymptotic habituation model (AHM) or the habituation-tedium model (HTM).

Finally, there was no evidence to conclude whether the participants preferred one type of ad saliency over another. It seemed that, since they were generally not consciously aware of the banner ads' existence (as evidenced by the memory tests results), the banner ads' salience level did not affect participants' attitude toward them.

In an effort to resolve the tension between the two models, a second study was deployed that focused on novice Internet users as the target participant. The reason being that the student population used for the first study were already

habituated to ads (the WUSE mean score was 85.02, SD=6.98). Therefore, any results that were obtained represented the tail end of each of the habitation models. To construct a reliable model of the effect of habituation on ad avoidance, one must *simulate* the effect using participants who have yet to be habituated to this type of stimuli.

In addition, since direct measures of attitude toward the banner ads failed to show a difference between saliency levels, covert measures—such as galvanic skin response and heart-rate—were employed to detect subtle differences in mental processing between high and low salient ads.

STUDY 2—NOVICE USERS

Experimental Design

The experiment design for the second study was nearly identical to that of the first. However, instead of three levels of ad saliency, I used only two levels: high and low (color animated vs. black and white non-animated). Therefore, the experiment design was a 2x2 mixed factorial using two independent variables:

Search Type (known-item vs. exploratory), which was manipulated within subjects, and Ad Saliency (high vs. low), which was manipulated between subject groups.

The primary dependent variable was orienting response to the banner ads and was measured using eye fixations, galvanic skin response (GSR), and heart rate.

Participants

The inclusion criteria for this study were as follows:

- 1. Ability to read and communicate in English.
- 2. 18 years old and above.
- 3. Unaided vision (do not wear glasses).
- 4. Novice Internet users as defined by scoring under 60 points on the WUSE.

One of this study's greatest challenges was recruiting novice Internet users.

Because children and the elderly cannot be recruited due to ethical and visual acuity concerns, I was limited to a small pool of potential subjects. To recruit the participants, I sent a mass email communication to the all the students and

employees of the University of North Carolina at Chapel Hill, posted flyers on campus and at local public libraries, and visited local computer literacy classes. Those who were interested to participate in the study were asked to call a toll-free number and answer an automated WUSE version using the phone's keypad. In addition they were asked to leave their name and number. Potential study participants who were qualified based on the inclusion criteria were called back within 48 hours to schedule their session. In the end, I was able to recruit only five participants. One participant was dropped due to low cognitive skills. In total, the results in this study were obtained from 4 participants (3 males, one female, mean age of 50.75).

In exchange for participating in this study, each participant received \$30 plus reimbursement for two-hour parking. Early termination would have resulted in \$20 of compensation.

Materials

Hardware

The workstation and eye tracker hardware was identical to the previous study. The GSR and heart rate data were collected using a Biopac MP36/35 Data Acquisition Unit that was hooked up to the participants via wired electrodes to the index and middle fingers of the participants' dominant hand (all the participants were right-handed), left ankle, and right forearm. The beginning and end times of the participants' interaction with the banner ads while reading the website articles were manually marked on the digital data strip using a hand switch, operated by the experimenter. The hand switch was pressed once at the start of the trial and twice

to signal trial completion. The signal obtained from the hand switch was synched with the rest of the data; this allowed the selection of only relevant data strips throughout the data analysis (see Figure 44).

Stimuli

The website, its content, the tasks, and the banner ads were identical to the previous study.

Procedure

As with the previous study, the experiment took place at the iDL. Upon arrival, the participants were seated approximately 50 centimeters from the workstation display and were asked to read the information sheet (Appendix H), and sign the consent form (Appendix I). The participants were informed of the experiment goals and procedure and were encouraged to ask questions. Then, the participants were hooked-up to the Biopac machine to start recording the GSR and heart rate data. After that, the experimenter began calibrating the eye tracker.

Before beginning the trials, the experimenter showed all the participants how to use the computer mouse and navigate the website. Once the participants understood the procedure, they were instructed to begin with the first trial. Unlike the previous study, there were no breaks during the data collection.

Results

Given the small sample size, the analysis was limited to the effect of repetition on the orienting response to banner ads. The effect of search type and ad saliency were not included.

Gaze Data

The eye tracking data was collected and analyzed using an identical mechanism as described in the Study 1 results section.

As with the previous study's analysis, this study's trials were consolidated into five samples. Using repeated measures ANOVA, there was no significant effect of repetition on number of fixations: $F_{(4,12)}=1.97$, ns., $partial\ \eta^2=.396$. However, using a Curve Estimation procedure, there was significant fit to inverse curves; see Table and Figure 44.

Table 7

Model fit results for the effect of repetition on number of fixations based on novice users.

Curve	R Square	F	df1	df2	Sig.
Inverse	.259	6.31	1	18	0.02
Quadratic	.272	3.18	2	17	0.67

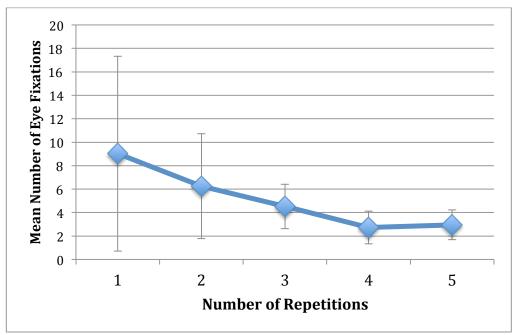


Figure 44. Distribution of the fixation data as a function of repetition among novice users.

GSR and Heart Rate

The data parsing relied on time marks recorded on the data strip (see Figure 45). However, because these marks were made manually, there was a large discrepancy between the number of recorded trials and sub-trials 14 from one participant to another. Since the number of trials for all the participants was a denominator of 5^{15} , the data was consolidated to five blocks for each participant.

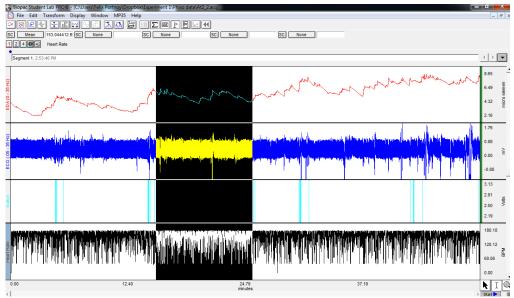


Figure 45. Depiction of the data analysis using Biopac Student PRO v.3.7.3.

Using repeated measure analysis, I found a significant main effect for orienting response as a function of repetition: $F_{(4,12)}=8.2, p\leq 0.01, partial\ \eta^2=.732$. Contrasts analysis revealed that the source of the effect lay in a *positive* linear distribution of the GSR data: $F_{(1,3)}=9.32, p=0.55, partial\ \eta^2=.757$ (see Figure 46).

-

¹⁴ Single web article.

¹⁵ Participant 1: 10 trials; Participant 2: 15 trials; Participant 3: 35 trials; Participant 4: 25 trials.

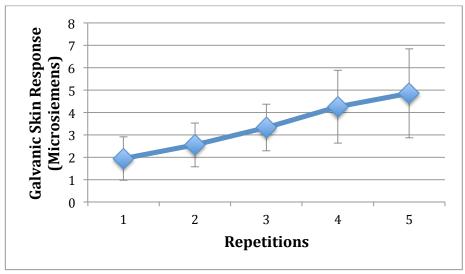


Figure 46. Distribution of the galvanic skin response data as a function of repetition.

Using the same analysis method, I found no significant effect as measured by heart rate: $F_{(4,12)}=0.19, ns.$, partial $\eta^2=.059$ (see Figure 47).

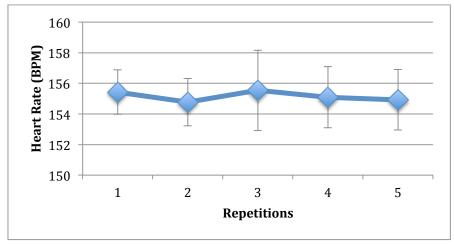


Figure 47. Distribution of the heart rate data as a function of repetition.

Summary

The results obtained from the GSR data seem to stand in contrast with both the AHM and HTM predictions. Since the participants were all novice users, the increase in the GSR is likely a result of increased frustration with the apparatus rather than a response to the banner ads; this would confound the response to the banner ads with the response to the computer interface. The unusually high heart

rate data (normal BPM among adults is 60-80) suggested a potential error during data collection. Taken together, the GSR and heart rate data posed face validity issues with regards to measuring physiological response to banner ads.

In contrast, the eye tracking data supported the asymptotic habituation model (AHM). Given the small sample size (N=4), it is not surprising that I did not receive a significant main effect of repetition. However, the large effect size and the results obtained from the curve fit analysis point to a Type II error. Additional data should be collected to validate this assertion.

DISCUSSION

Search Type Matters

The findings presented in this dissertation provide strong support to the search type hypothesis: participants engaged in a known-item search task were *less* likely to look at banner ads as compared with participants engaged in an exploratory search task. As expected, when engaged in known-item search, participants focused on specific keywords to find their target. Therefore, as a result of their high perceptual load, they were not able to process additional perceptual distractors, which made them less likely to look at the banner ads (Lavie et al., 2004). As shown in Figure 48, rather than focus on target keywords for a known-item search, exploratory searchers distributed their attention across the entire webpage to synthesize as much information as possible. Since they were continuously prioritizing the information presented in the article and relating it to their search task, their short-term memory reached its full capacity. As a result, their ability to block distractors was diminished, leading to an increased number of fixations on the banner ads.

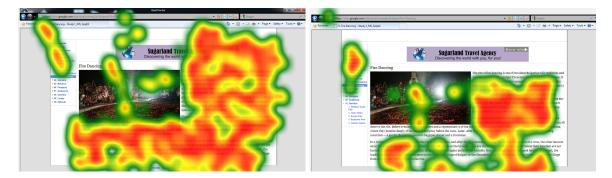


Figure 48. Heat map showing the fixation distribution in exploratory search task (left image) and known-item search task (right image).

One reason previous studies reported conflicting results in the performance of known-item vs. exploratory search tasks was that the search task definition varied greatly from one study to another. In some cases, exploratory or "experiential" search was defined variously as "aimless browsing" or a series of known-item search queries. In those cases, the variable definition of an exploratory search led to erroneous interpretations of the search type effect. Therefore, a crucial challenge for the research community would be to reach a consensus as to the basic definition of various search activities and their operationalization.

Another contributor to previous inconsistent findings may have been the type of measure used to capture the orienting response to banner ads. As cited in the literature review, the present CTR for banner ads is very small (less than 1%), indicating that most users pay little or no attention to them. As evidenced by the memory results in study 1, attempts to measure orienting response through the users' conscientious response—either by mouse clicks or memory tests—may likely be hindered by a floor effect. To overcome the floor effect, previous studies that relied on memory as their primary measure had to create a unique website structure that artificially enhanced the appearance of the ad (see Figure 5 on page 6

and Figure 6 on page 7) or substantially limited the ad variation (Calisir & Karaali, 2008). In both of these cases, the ecological validity was diminished since the findings were less likely to be replicated in the real world.

The studies described in this dissertation approached the problem by explicitly attempting to simulate real-world websites as much as possible, given the constraints of the experimental design. In addition, eye tracking proved to be an indispensible tool to measure the participants' orienting response. Although the participants did not retain much of the information they observed, this did not diminish the validity of the findings. Most ad campaigns run for weeks at a time to build up the brand awareness; therefore, given the short interaction that my participants had in the lab with multiple mock brand names, it is not surprising that the participants failed to retain the content information of the banner ads.

Less is More

The second hypothesis predicted a higher orienting response toward low salient banner ads as compared with medium and high salient banners ads. The results presented here showed that high salient ads were less likely to attract the participants' attention as compared with low salient banner ads. Specifically, animated color ads were more likely to be identified as distractors as they were markedly different from the textual content that the participants were asked to review. Regardless of the participants' type of search task, animated ads were given less attention than black-and-white, non-animated ads.

While this finding coincided with other results obtained from different research domains (Folk et al., 1992; Pashler, 1999; Tullis et al., 2009; Van Orden et

al., 1993), this study marked the first appearance of this effect in the online advertising domain. The results obtained here stand in contrast to previous findings from the print advertising domain (Barlow & Wogalter, 1993; Berdie, 1992; Lohse, 1997; Lutz & Lutz, 1977; Valiente, 1973), which may suggest different perceptions of web content vs. print media.

Earlier findings by Nielsen (1997) showed that participants tended to *scan* web content rather than read it word-for-word. This may also explain the results obtained by a number of studies that found that people are less likely to retain promotional and informational content that was obtained from web sources as compared with print-format (Eveland & Dunwoody, 2002; Jones, Pentecost, & Requena, 2005; Sundar, Narayan, Obregon, & Uppal, 1998). These results would seem to indicate that web users may be less engaged with online content, which could make them more effective in inhibiting salient distractors as compared with print media.

The current study's questionnaire results measuring the participants' attitude to the ads showed no significant effect on the user experience. In fact, based on the results, the participants showed *indifference* to the banner ads' saliency level (the mean range between the ad salience levels ranged from 2.77 to 3.22 on a 7-point scale). Perhaps a non-intrusive placement for the banner ad (see Figure 34 on page 81) made it easier to ignore the ad regardless of its saliency level. If the participants had been forced to confront the banner ad—for example, by embedding it in the text—the results could have showed a preference for one level of saliency over another.

The Effect of Habitation on the Response to Banner Ads

A question that this study attempted to answer was how or whether habituation affected the users' response to banner ads. Since there have been no published attempts to answer this question in the online advertising domain, two leading models were considered: the asymptotic habituation model (AHM) and the habituation-tedium model (HTM). Because the first study did not yield conclusive result on this issue with a web-savvy student population, the second study attempted to resolve this question using novice web users. The assumption was that, because students are presumably already habituated to banner ads, their response reflected only the tail end of the habituation effect. Despite the small number of participants (N=4), I did find preliminary support to AHM using the eye gaze data. Although the main effect was not significant, the relatively large effect size ($partial \eta^2 = .396$) along with the curve analyses suggested that the users' response to banner ads followed an asymptotic (or "inverse") curve. However, additional data would need to be collected to validate this assertion.

The Use of Physiological Measures

One of the second study's goals was to evaluate the benefits of galvanic skin response (GSR) and heart rate as measures of orienting response to banner ads.

Unlike any of the AHM or HTM predictions, the result from the GSR measure showed a strong positive linear effect of repetition on the orienting response to the banner ads. Despite the small sample size, this result was puzzling. The most plausible explanation was that GSR measured the participants' discomfort level rather than

their reaction to the banner ads. Since the participants were all novice users, in many cases they struggled both to navigate the website and to use the computer mouse. By using novice users, I may have introduced confound into the experimental design. The participants' heart rate data showed an unusually high heart rate, which may be a result of erroneous calibration of the equipment.

Overall, the physiological data that was collected in the second study was not suitable for the purpose it aimed to achieve. While there have been reports that these measures successfully captured the orienting response to banner ads (e.g., Lang et al., 2002), those studies presented the banner ads in complete isolation (only the banner ad was shown on the screen) and so substantially compromised the results' ecological validity.

Recommendations Based on the Research Findings

Based on the study results described in this dissertation, I would suggest the following recommendations to web publishers who want to improve CTR while preserving a satisfying user experience:

Banner ads should be displayed in low salience to attract the users' attention. Specifically, the appearance of the ad should resemble the graphical appearance of the webpage. It is important to note that saliency, as defined earlier in this dissertation, is not a fixed trait. The level of saliency of the ad will rely on its resemblance (or lack of) to its surrounding objects (see Figure 49).

Low salient banner ad Highly salient banner ad

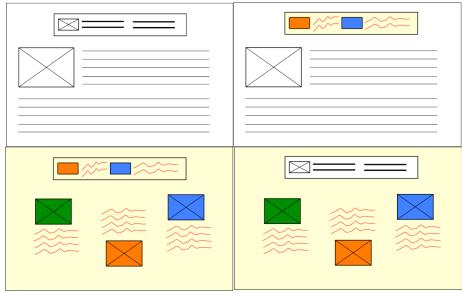


Figure 49. An illustration of the differing salience levels of the banner ads as a function of their surrounding objects.

Most modern website graphic analyzers are designed to detect pornography (see a review by Wen, Zhu, & Peng, 2009). However, this technology can be converted by online advertising exchange platforms to focus instead on the analysis of webpage color and contrast distribution. Therefore, in addition to matching keywords with relevant ads, designers could match a graphical variant of the banner ad to the host webpage to ensure low saliency.

In addition to paying web publishers by the number of clicks the ad generated, another popular method is cost per thousand impressions (CPT). With CPT, advertisers pay the publishers for providing the web space to place their ad. Each time the ad is placed it is counted as a single impression. Today, the advertising industry does not differentiate the cost of impressions; rather, the publishers charge a flat rate. However, based on this study's results, web users may be more likely to notice ads during an exploratory search activity as compared with

a known-item search. Therefore, the value of the impression during an exploratory search could be much greater and priced accordingly.

At present, efforts to automate query intent focus on the classification of informational, navigational, and transactional (with emphasis on commercial intent) queries (Ashkan, Clarke, Agichtein, & Guo, 2009; Lewandowski, Drechsler, & von Mach, in press) with reported accuracy rate of 74% (Jansen, Booth, & Spink, 2008). While some argue that a navigational query is a form of known-item search (Broder, 2002), there have not been published attempt to differentiate informational queries between known-item and exploratory search types. Hopefully, the results presented here will encourage additional research in this field.

Although the findings that support the AHM are not conclusive, publishers should consider changing the rate of CPT as a function of repetition, such that the cost of the first few thousands of impressions will be greater than those that follow. This could also encourage advertisers to prolong their campaigns since the cost would be substantially diminished over time.

While the focus of the studies described in this dissertation has been on banner ads, the findings are applicable to any promotional or low-priority informational messages that may distract the user from their primary task. For example, an online bank that introduced a new design to their ATM receipts should present the new design using low salience. As shown in Figure 50, the banner ad uses a color scheme similar to that used in the webpage's content area. Since the primary goal of the user is to review their financial transaction, a low salient banner

might allow the promotional message to be visible without imposing itself on the user.

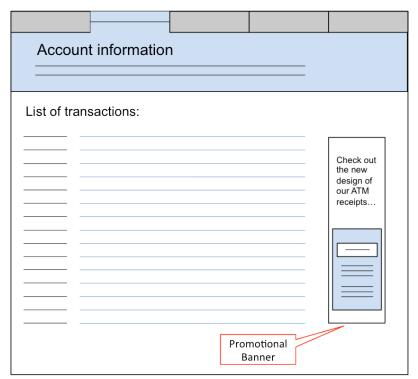


Figure 50. An illustration of a low salient promotional message embedded in a financial webpage.

FUTURE WORK

Studying the responses of novice web users proved to be an effective way to simulate early exposure to banner ads. However, further work needs to be done to confirm the assertion that the AHM best describes the users' response to banner ads as a function of repetition. Since the recruitment of novice web users in this day and age is quite challenging, other alternatives should be considered. For example, instead of inviting participants to a lab, the research could be done in the field using a portable eye-tracking device, such as the Tobii X1 Light Eye Tracker (see Figure 51).

Another alternative would be to monitor an experienced user population, such as students, using a longitudinal log study. For example, one could introduce a promotional message in the school's website and monitor the CTR over a period of several weeks or months. In addition



Figure 51. The Tobii X1 Light Eye Tracker allows the researcher to conduct eye-tracking studies in the field using a laptop.

to CTR, mouse movements have been shown to highly correlate with eye movement (Chen, Anderson, & Sohn, 2001; Rodden, Fu, Aula, & Spiro, 2008). Therefore, tracking the users' mouse movements over time (using log analysis) could also reveal users' interaction pattern with banner ads.

Still unresolved is whether a banner ad's salience affects the user experience. The current studies showed no difference in the participants' attitude between the various saliency levels of banner ads. These results echoed a previous finding by Burke et al. (2005) who also reported finding no difference in mental workload between animated vs. non-animated commercial banners. However, for both Burke et al. (2005) and the current studies, the results may be due to the minimal mental demand that the primary task required. Had the tasks been more challenging—for example, by introducing a time constraint and/or embedding the ad in the content area—the impact of high salient banner ads on the user experience might have been greater.

The current studies' findings suggest that, under normal circumstances, web users may have grown accustomed to various levels of ad saliency. While this may be true with desktop computer-based browser interfaces, previous studies with mobile interfaces found an increase in the level of intrusiveness caused by ads (Heinonen & Strandvik, 2007; Tsang, Ho, & Liang, 2004). Given the small display sizes of mobile device screens, it would be interesting to investigate whether the effect of ad saliency that was reported in this dissertation would be replicated on mobile devices.

While this study manipulated saliency levels, it was investigated only with one type of highly salient ad—color animation. In addition, the use of a color image across all the web articles may have confounded the salience manipulation of the banner ads. Although the black and white banner ad was designed to have low salience with relation to the textual webpage, having a color image directly

underneath it, may have increased its salience level. To increase the validity of my findings, future studies should replicate the results obtained here with a reverse set of ad salience levels using a content area that includes animated colorful images with a content image that matches the salience level of the banners.

Finally, the search type hypothesis derived its prediction based on the assumption that exploratory search leads to a high cognitive load and known-item search leads to high perceptual load. However, mental load was not directly measured in this study. Despite the alignment of the results obtained here with the load theory predictions (Lavie et al., 2004), there has not been an empirical support to this assumption.

CONCLUSIONS

Earlier in this dissertation I presented this core research question: "which factors affect users' perception of online banner advertisements?" Through an empirical investigation employing both experienced and novice web users I have demonstrated the following:

- Web users respond differently to online banner ads as a function of their search behavior. Specifically, users are less likely to look at banner ads while engaged in a known-item search as compared with an exploratory search.
- Web users are less likely to look at banner ads that are characterized by high salience as compared with low salience.
- 3. There is some evidence that suggests a decrease in attention to banner ads as a function of their repetition.

These findings were translated into a set of design recommendation intended for both online publishers and advertisers. If these recommendations were to be correctly implemented, then perhaps more revenue could be generated from online banner ads.

Although the online advertising industry's current focus is on increasing behavioral targeting, the guidelines presented here do not require users to give up their privacy in exchange for free access to web content. By analyzing the structure of a user's search query and the ad's host webpage, ads could be placed such that they would be more effective *without* compromising the user experience. Thus,

users and publishers could reach a win-win solution where publishers could sustain their business model while preserving free access for the public.

APPENDICES

Appendix A – Search Tasks

Known-Item Search Tasks¹⁶

South Korea

- 1. How many pieces of historic relics were excavated from Anapji Pond?
 - a. 9000
 - b. 18000
 - c. 1700
 - d. 33000
- 2. How many square kilometers does the Seoraksan National Park cover?
 - a. 327.4
 - b. 13.9
 - c. 470.1
 - d. 163.6
- 3. To which city the National Museum of Korea was relocated?
 - a. Yongsan-gu
 - b. Cheongiu
 - c. Incheon
 - d. Seongnam
- 4. In which year did King Taejong move to Hanyang?
 - a. 1405
 - b. 535
 - c. 1741
 - d. 236

¹⁶ The correct answers are highlighted.

- 5. How many malls are located in Dongdaemun Market? a. 7 b. 46 c. 26 d. 43 Bulgaria 1. On the foundations of which basilica the archaeological museum was built? a. Basilica of the Holy Mother of God Eleusa b. Basilica of the Merciful Virgin c. Basilica of Pliska d. Basilica of Bishop Yoan 2. What is Olimpi Panov famous for? a. Established the central bank b. First president of Bulgarua c. Liberated Bulgaria from the Ottoman rule d. Signed a truce with Prussia
 - 3. What is the name of the town where the Museum of Wood-Carving is located?
 - a. Kazanluk
 - b. Peshtera
 - c. Tryavna
 - d. Velingrad
 - 4. In which year was the Sofia Synagogue partially destroyed as a result of bombings?
 - a. 1916
 - b. 1944
 - c. 1943
 - d. 1932
 - 5. In which date does the Fire Dance take place?
 - a. June 3rd
 - b. July 28th
 - c. March 11th
 - d. September 19th

Slovakia

- 1. Which queen restored the significance of Bratislava?
 - a. Queen Maria Theresa
 - b. Queen Mary
 - c. Queen Elizabeth II
 - d. Queen Teuta
- 2. In which year the Slovak National Museum was established?
 - a. 1903
 - b. 1961
 - c. 1973
 - d. 1993
- 3. What is the name of the architect who designed the National House?
 - a. Ilija Arnautović
 - b. Matevž Čelik
 - c. Emil Beluš
 - d. Niko Kralj
- 4. Which part in the Orava Castle is open for the public?
 - a. Servants' Quarter
 - b. Gate Hoise
 - c. Tower
 - d. Dungeon
- 5. How many pools can be found in Tatralandia?
 - a. 7
 - b. 12
 - c. 25
 - d. 11

Moldova

- 1. What is the last name of the architect who designed the National Museum of Ethnography and Natural History?
 - a. Tsigankov
 - b. Galchinski
 - c. Meldaze
 - d. Kara

a. 1979 b. 1783 c. 1935 d. 1964 4. Which structure is located above the entrance to Soroca Fort? a. Visitor Center b. Wine Celler c. Guard Post d. Military Church 5. What is the distance between Chisinau and Serpeni monument? a. 39 km b. 55 km c. 85 km d. 40 km Paraguay 1. Which collected in the Museo del Barro is most popular? a. Mayan dresses b. Balam arrowheads c. Guarani ceramics d. Votan paintings 2. What was the military rank of Mariano Recalde? a. Captain b. Lieutenant c. General

3. Before reopening in 1990, what year was the Saharna Monastery closed?

2. In what depth the bottles of wine are store in Milestii Mici?

a. 80 metersb. 55 metersc. 90 metersd. 140 meters

d. Colonel

3.	How many waterfalls are in Iguassu Falls?
	a. 275
	b. 55 c. 215
	d. 116
4.	What is the meaning of the logo that Moses Bertoni used in his forms?
	a. Humble immigrant
	b. Paraguay is mec. From the jungle
	d. Victory and justice
5.	How many years did it tale to build the Holy Trinity of Paraná?
	a. 61
	b. 97 c. 194
	d. 121
Guatem	nala
1.	Which airport was affected by the 1998 eruption of Pacaya Volcano?
	a. Chiquimula
	b. La Aurorac. Mundo Maya
	d. Puerto Barrios
2.	What was the type of entertainer seen at Parque Central?
	a. Story Teller
	b. Jugglerc. Clown
	d. Musician
3.	How many archeological exhibits are displayed at the Museo Popol Vuh?
	a. 800
	b. 400c. 700
	d. 500

- 4. How many square miles of jungle comprise Tikal National Park?
 a. 222
 b. 286
 c. 64
 d. 383
- 5. Which doors should tourists use when they enter the church in ChiChi Market?
 - a. Yellow doors
 - b. Back doors
 - c. White doors
 - d. Side doors

Namibia

- 1. What is the name of the ocean current that affects the temperature of the Skeleton Coast Park?
 - a. Mocambique
 - b. Oyashio
 - c. Canary
 - d. Benguela
- 2. How many bird species have been recorded at The Daan Viljoen Game

Reserve?

- a. 200
- b. 150
- c. 240
- d. 120
- 3. What does the name "Etosha" mean?
 - a. Salt land
 - b. Place of worship
 - c. Bushland
 - d. Place of dry water

- 4. What is the name of the minority group of people living in Bwabwata National Park?
 - a. Lozian
 - b. Caprivian
 - c. Subian
 - d. Mbukushuan
- 5. Which region in Namibia has example of Bushman rock art?
 - a. Erongo
 - b. Damaraland
 - c. Ohangwena
 - d. Kaokoland

Turkey

- 1. Which ruler built the first building of Hagia Sophia church?
 - a. Maximinus I
 - b. Gordian I
 - c. Constantius I
 - d. Licinius I
- 2. What does the name "Pamukkale" mean?
 - a. Pool of sapphire
 - b. Majestic view
 - c. Cotton castle
 - d. Water of the earth
- 3. What is the cost of admission to the terrace houses?
 - a. € 14.15
 - b. € 2.85
 - c. € 8.35
 - d. € 10.15
- 4. What is one of the flower patterns used in the tiles covering the Blue Mosque?
 - a. Banksia
 - b. Cornflower
 - c. Gladiolus
 - d. Tulip

- 5. Whom did the Attalids fight against during the first and second Macedonian Wars?
 - a. Philip V of Macedon
 - b. Hannibal of Carthage
 - c. Nabis of Sparta
 - d. Pandion II of Athens

Djibouti

- 1. What is Djibouti's national pastime?
 - a. Playing Backgammon
 - b. Chewing Qat
 - c. Hunting Jerboa
 - d. Watching a local TV show
- 2. Which species flourishes in Lake Assal?
 - a. Bacteria
 - b. Lizards
 - c. Insects
 - d. Scorpions
- 3. Which foreign Djibouti residents favor Khor Ambado?
 - a. German
 - b. English
 - c. French
 - d. Dutch
- 4. For which type of bird Lake Abbe is known?
 - a. Imperial Eagles
 - b. Storm-Petrels
 - c. Pelicans
 - d. Flamingos
- 5. In the 19th century, how many Ethiopian slaves were shipped via Tadjoura per year?
 - a. 8000
 - b. 12000
 - c. 6000
 - d. 4000

Bhutan

- 1. Which king the National Memorial Chorten honors?
 - a. Tenzin Drukgye
 - b. Jigme Dorji Wangchuck
 - c. Mipham Wangpo
 - d. Chhogyel Sonam Gyaltshen
- 2. How many temples did King Songtsen Gampo build?
 - a. 61
 - b. 263
 - c. 108
 - d. 141
- 3. By which name Gyelchok's descendants are known for?
 - a. Children of Parob
 - b. Lords of Humrel
 - c. Singye Eagles
 - d. Drugyel Sun-rays
- 4. What does the word "Dochey" mean?
 - a. Leader
 - b. Courtyard
 - c. River
 - d. Mansion
- 5. Which animal is being referred to as a "bee-stung moose"?
 - a. Seadragon
 - b. Sun Bear
 - c. Aye-aye
 - d. Takin

Exploratory Search Tasks

- 1. **Bhutan:** The Bhutanese measure their quality of life using the "gross domestic happiness." Based on the articles you read, please postulate why they chose this measure as oppose to the standard income measure. Does this measure describe their quality of life more accurately?
- 2. **South Korea:** Does the natural surrounding affect the traditional culture in South Korea? Please explain your answer.
- 3. **Bulgaria:** Bulgarian culture has been influenced by many religious traditions. Based on the articles you read, how do modern Bulgarians integrate these traditions in the current day?
- 4. **Slovakia:** What is your perceived condition of the infrastructure in Slovakia? Does that affect your interest to travel in this country? Please explain.
- 5. **Moldova:** Since the collapse of the Soviet Union, Moldova has been making attempts to restore its national heritage. Based on the articles your reviewed, were they successful? Please explain.
- 6. **Paraguay:** What effect (if any) did the European settlers had on Paraguay? Please refer to the articles in your response.
- 7. **Guatemala:** What is the relationship between descendants of the Spanish settlers and the indigenous people of Guatemala?
- 8. **Namibia:** Does the Namibian government allocate sufficient resources to protect its natural and wildlife resources? Please explain your answer based on the articles you read.
- 9. **Turkey:** What was the Ottoman empire's attitude towards the heritage left by the former Roman and Greek rulers of Turkey?
- 10. **Djibouti:** How safe it is to travel in Djibouti? Please explain your rationale.

Appendix B - Information Sheet

Information sheet

University of North Carolina-Chapel Hill

Information about a Research Study

IRB Study: 11-2232

Consent Form Version Date: 10-11-2011
Title of Study: The Perception of Online Content
Principal Investigator: Felix Portnoy, PhD Student

Faculty advisor: Gary Marchionini

UNC-Chapel Hill Department: School of Information & Library Science

UNC-Chapel Hill Phone number: 919.966.3611

Email Address: march@ils.unc.edu

Study contact telephone number: 269.767.8669 Study contact email: portnoy@email.unc.edu

What are some general things you should know about research studies?

You are being asked to take part in a research study. To join the study is voluntary. You may refuse to join, or you may withdraw your consent to be in the study, for any reason, without penalty.

Research studies are designed to obtain new knowledge. This new information may help people in the future. You may not receive any direct benefit from being in the research study (other than the financial compensation). There are no foreseeable risks to being in this research study.

Details about this study are discussed below. It is important that you understand this information so that you can make an informed choice about being in this research study. Once you have read this informational sheet, you will be asked to fill in a consent form to agree to participate in this study. You should keep this informational sheet for your records. If you are unwilling to participate after reading the informational sheet, then you are free to leave the study without penalty. In fact, you are free to leave the study at any time should you decide to withdraw your consent.

You should ask the researcher if you have questions about this study.

What is the purpose of this study?

The purpose of this research study is to understand how people perceive information in the World Wide Web. In addition, we are interested to explore how online content affects our cognitive systems.

How many people will take part in this study?

This form has been approved by the IRB for use between and If you decide to be in this study, you will be one of approximately 60 participants in this research study.

How long will your participation in this study last?

Your participation in this study will take approximately 1.5 hours.

What will happen if you take part in the study?

This study will ask you to perform several search tasks based on pre-loaded search results. You will be asked to read travel information about foreign countries and answer specific questions based on the information that you will read. While you perform your tasks, the system will gather information pertaining your eye movements. In addition, in the end of the study, you will be asked to complete several questionnaires.

What are the possible benefits from being in this study?

This study may not benefit you directly (other than the financial compensation), but your participation will help us understand how users interact with online information and help to derive better interface designs to support this task while maintaining a positive user experience.

What are the possible risks or discomforts involved from being in this study?

There are no foreseeable risks to being in this study. However, there may be uncommon or previously unknown risks. You should report any problems to the researcher. Please use the email address or phone number provided if problems arise after you have completed participation.

How will your privacy be protected?

You will not be identified in any report or publication about this study. After you have completed participation, we will destroy all links between your personal information (e.g., name and email address) and your data. This information will only be used to schedule your experimental session. Once you have completed the study, there will be no written or electronic record that you participated in this study. Data will only be identifiable by username, and there will be no record linking you to the username you used during the study.

All data that you generate while participating in the study will be stored on a secure server in a password-protected account and will only be available to the researchers.

Will you receive anything for being in this study?

Yes. You will be given \$15 for your participation. Early termination will result in \$10 compensation.

Will it cost you anything to be in this study?

It will cost you nothing to be in this study, other than time to come to the UNC campus location and participate in the study.

What if you are a UNC student?

You may choose not to be in the study or to stop being in the study before it is over at any time. This will not affect your class standing or grades at UNC-Chapel Hill. You will not be offered or receive any special consideration if you take part in this research.

What if you have questions about this study?

You have the right to ask, and have answered, any questions you may have about this research. If you have questions, or concerns, you should contact the researchers listed on the first page of this form.

What if you have questions about your rights as a research participant?

All research on human volunteers is reviewed by a committee that works to protect your rights and welfare. If you have questions or concerns about your rights as a research subject you may contact, anonymously if you wish, the Institutional Review Board at 919-966-3113 or by email to IRB_subjects@unc.edu.

Appendix C – Consent Form

Title of Study: The Perception of Online Content

IRB Study Number: 11-2232

Consent Form Version Date: 10-11-2011 **Principal Investigator:** Felix Portnoy

Participant's Agreement:

I have read the information provided the questions I have at this time. I vo study.	
Signature of Research Participant	 Date

Printed Name of Research Participant

Signature of Person Obtaining Consent Date

Felix Portnoy

Printed Name of Person Obtaining Consent

Appendix D – Recall Questionnaire

This section examines your memory for the banner ads from the articles you just reviewed. Please answer each question as best as you can. If you do not remember the answer, go ahead and guess.

The following questions pertain the ad about a TRAVEL AGENCY:

- 1. Name the company that appeared in the ad:
- 2. What object was included in the banner ad?
- 3. What was the slogan of the ad?

The following questions pertain the ad about an AIR LINE:

- 1. Name the company that appeared in the ad:
- 2. What object was included in the banner ad?
- 3. What was the slogan of the ad?

The following questions pertain the ad about CAR RENTAL:

- 1. Name the company that appeared in the ad:
- 2. What object was included in the banner ad?
- 3. What was the slogan of the ad?

The following questions pertain the ad about HOTELS:

- 1. Name the company that appeared in the ad:
- 2. What object was included in the banner ad?
- 3. What was the slogan of the ad?

The following questions pertain the ad about RESORTS:

- 1. Name the company that appeared in the ad:
- 2. What object was included in the banner ad?
- 3. What was the slogan of the ad?

Appendix E – Recognize Questionnaire

This section examines your memory for the banner ads from the articles you just reviewed. Please answer each question as best as you can. If you do not remember the answer, go ahead and guess.

The following questions pertain the ad about a TRAVEL AGENCY:

- 1. Name the company that appeared in the ad:
 - a. Kalos
 - b. Maupin
 - c. Sugarland
 - d. Travelink
- 2. What object was included in the banner ad?
 - a. Map
 - b. Tower
 - c. Suitcase
 - d. Globe
- 3. What was the slogan of the ad?
 - a. Great places to go and see, that's what we guarantee...
 - b. Miles and miles of smiles
 - c. Discovering the world with you, for you!
 - d. Travel with the PROS, Just say when, and leave the rest to us!

The following questions pertain the ad about an AIR LINE:

- 1. Name the company that appeared in the ad:
 - a. Bosna
 - b. Omni
 - c. Engiadina
 - d. Atlas
- 2. What object was included in the banner ad?
 - a. Bird
 - b. Star
 - c. Seats
 - d. Deer

- 3. What was the slogan of the ad?
 - a. Something special in the air
 - b. We love to fly and it shows
 - c. Making the sky the best place on Earth
 - d. Sit back and relax

The following questions pertain the ad about CAR RENTAL:

- 1. Name the company that appeared in the ad:
 - a. Juno
 - b. Eldan
 - c. VIP
 - d. Tamir
- 2. What object was included in the banner ad?
 - a. Key
 - b. Road
 - c. Headlights
 - d. Bull
- 3. What was the slogan of the ad?
 - a. Rent your car today!
 - b. Love your road.
 - c. You rent more than a car, you rent our service!
 - d. Pick us and we'll pick you up.

The following questions pertain the ad about HOTELS:

- 1. Name the company that appeared in the ad:
 - a. Marion
 - b. Abigail
 - c. Wild Boar
 - d. Lato
- 2. What object was included in the banner ad?
 - a. Crown
 - b. Lion
 - c. Flower
 - d. Bell

- 3. What was the slogan of the ad?
 - a. Twice the comfort, twice the value, twice the Hotel
 - b. It's not a hotel, it's a way of life
 - c. Distinctive service for the busy traveler
 - d. The art of meeting your highest expectations

The following questions pertain the ad about RESORTS:

- 1. Name the company that appeared in the ad:
 - a. El San
 - b. Shoreline
 - c. Tempah
 - d. Palmilla
- 2. What object was included in the banner ad?
 - a. Hammock
 - b. Pool
 - c. Sunglasses
 - d. Sunset
- 3. What was the slogan of the ad?
 - a. How about some peace and quite?
 - b. Who's taking care of you?
 - c. The right way, your way
 - d. Where leisure meets luxury

Appendix F - Attitude Towards Ads Questionnaire

This section pertains to an overall evaluation of the ads based on the following scales. Notice that some of the scales are reversed. So please read both ends of the scale carefully before making your choice.

Appeal Appealing	1	2	3	4	5	6	7	Unappealing
Informativene	ess							
Informative	1	2	3	4	5	6	7	Non- Informative
Excitement Unexciting	1	2	3	4	5	6	7	Exciting
Interest Boring	1	2	3	4	5	6	7	Interesting
Goodness Good	1	2	3	4	5	6	7	Bad
Pleasantness Pleasant	1	2	3	4	5	6	7	Unpleasant
Tedium Dull	1	2	3	4	5	6	7	Dynamic
Confusion Clear	1	2	3	4	5	6	7	Confusing
Attractiveness Unattractive	s 1	2	3	4	5	6	7	Attractive
Favorable Favorable	1	2	3	4	5	6	7	Unfavorable
Likable Likable	1	2	3	4	5	6	7	Dislikable
Sophistication Ordinary	1	2	3	4	5	6	7	Sophisticated

Persuasiven Persuasive	ess 1	2	3	4	5	6	7	Unpersuasive
Quality Low Quality	1	2	3	4	5	6	7	High Quality

Appendix G - Web User Self Efficacy Scale (WUSE)

WUSE was designed such that the higher the total score is, the more web experienced is the responder. However, to prevent Affirmation Bias, half of the statements are presented in a negative form (specifically, questions 3, 5, 6, 8, 9, 10, 13, 15, 16, 18). For example, if a participant responded '2' for question #3, than that answer should be converted to '4'.

Since the range of WUSE is 20 to 100, the threshold for novice users (as opposed to intermediate or expert) was selected to be 50.

1. I would not have any problems creating a simple web page.

Strongly	1	2	3	4	5	Strongly
Disagree	1	L	3	4	J	agree
2. I find usi	ing ema	il easy				
Strongly	1	2	2	4	-	Strongly
Disagree	1	۷	3	4	5	agree
3. I am not	really s	ure what	a browse	er is		
Strongly	1	2	3	1.	ς	Strongly

4. I know how to use software (e.g. Picasa, Photoshop) for editing and sharing images...

agree

Strongly 1 2 3 4 5 Strongly

Disagree

Disagree

agree

5. I would never try to download files from the Internet; that would be too complicated...

Strongly						Strongly
	1	2	3	4	5	
Disagree						agree

6. Using communication software, like MSN Messenger or Skype always cause me some problems...

Strongly	1	2	2	4	_	Strongly
	1	Z	3	4	5	
Disagree						agree

7. Finding my way around web sites is usually easy for me...

Strongly	1	2	3	4	5	Strongly
Disagree						agree

8. I much prefer using letters or the telephone to communicate with people, rather than the Internet...

Strongly						Strongly
	1	2	3	4	5	
Disagree						agree

9. I would not know how to capture pictures from the Internet...

Strongly						Strongly
	1	2	3	4	5	
Disagree						agree

10. Attaching a file to an email would be very difficult for me...

Strongly						Strongly
	1	2	3	4	5	
Disagree						agree

11. I know h	now to tes	st my com	puter for	the pres	ence o	of spyware				
Strongly	1	2	3	4	5	Strongly				
Disagree	1	L	3 1 3	3	agree					
12. Adding	12. Adding hypertext links is quite straightforward									
Strongly	1	2	3	4	5	Strongly				
Disagree	1	2	3	Т	5	agree				
13. If my co	mputer b	ecame in	fected wit	th a virus,	, I wou	ald not know how to get				
rid of it										
Strongly	1	2	3	4	5	Strongly				
Disagree	1	4	3	т	5	agree				
14. Using th	14. Using the Internet makes it much easier to keep in contact with people									
Strongly	1	2	3	4	5	Strongly				
Disagree	1	4	3	Т	5	agree				
15. Using so	cial netw	orks (e.g.	. Faceboo	k) to shar	e my	photos is too				
complicated	d for me									
Strongly	1	2	3	4	5	Strongly				
Disagree	1	L	3	4	3	agree				
16. I sometimes "get lost" when trying to navigate through the Internet										
Strongly	1	2	3	4	5	Strongly				
Disagree	1	4	J	7	5	agree				
17. I regularly exchange music and/or video files with friends										

Strongly Disagree	1	2	3	4	5	Strongly agree		
18. I have difficulty buying things from online stores (e.g. Amazon)								
Strongly	1	2	3	4	5	Strongly		
Disagree			3	4		agree		
19. I know how to deal with annoying advertisements that appear while I'm								
using the Internet								
Strongly	1	2	3	4	5	Strongly		
Disagree	1	4	3	1		agree		
20. I can usually sort out any Internet access problems I may encounter								
Strongly	1	2	3	4	5	Strongly		
Disagree	1	2	J	т		agree		

Appendix H – Information Sheet for Study 2

Information Sheet

University of North Carolina at Chapel Hill Consent to Participate in a Research Study Adult Participants

Consent Form Version Date: March 1st, 2012

IRB Study # 12-0255

Title of Study: Modeling the perception of online content

Principal Investigator: Felix Portnoy

Principal Investigator Department: School of Info & Libr Science

Principal Investigator Phone number: 269-767-8669
Principal Investigator Email Address: portnoy@unc.edu

Faculty Advisor: Gary Marchionini

Faculty Advisor Contact Information: gary@ils.unc.edu

What are some general things you should know about research studies?

You are being asked to take part in a research study. To join the study is voluntary. You may refuse to join, or you may withdraw your consent to be in the study, for any reason, without penalty.

Research studies are designed to obtain new knowledge. This new information may help people in the future. You may not receive any direct benefit from being in the research study (other than the financial compensation). There are no foreseeable risks to being in this research study.

Details about this study are discussed below. It is important that you understand this information so that you can make an informed choice about being in this research study. Once you have read this informational sheet, you will be asked to fill in a consent form to agree to participate in this study. You should keep this informational sheet for your records. If you are unwilling to participate after reading the informational sheet, then you are free to leave the study without penalty. In fact, you are free to leave the study at any time should you decide to withdraw your consent. You should ask the researcher if you have questions about this study.

What is the purpose of this study?

The purpose of this research study is to understand how people perceive information in the World Wide Web. In addition, we are interested to explore how online content affects our cognitive systems.

How many people will take part in this study?

If you decide to be in this study, you will be one of approximately 12 participants in this research study.

How long will your participation in this study last?

Your participation in this study will take approximately 1.5 hours.

What will happen if you take part in the study?

This study will ask you to perform several search tasks based on pre-loaded search results. You will be asked to read travel information about foreign countries and answer specific questions based on the information that you will read. While you perform your tasks, the system will gather information pertaining to your eye movements, skin conductivity, and heart rate. This will be done by attaching electrodes to your index and middle fingers in your dominant hand, in addition to your left ankle and right forearm. In addition, at the end of the study, you will be asked to complete several questionnaires.

What are the possible benefits from being in this study?

This study may not benefit you directly, but your participation will help us understand how users interact with online information and help to derive better interface designs to support this task while maintaining a positive user experience.

What are the possible risks or discomforts involved from being in this study?

There are no foreseeable risks to being in this study. However, there may be uncommon or previously unknown risks. You should report any problems to the researcher. Please use the email address or phone number provided if problems arise after you have completed participation.

How will your privacy be protected?

You will not be identified in any report or publication about this study. After you have completed participation, we will destroy all links between your personal information (e.g., name and email address) and your data. This information will only be used to schedule your experimental session. Once you have completed the study, there will be no written or electronic record that you participated in this study. Data will only be identifiable by username, and there will be no record linking you to the username you used during the study. All data that you generate while participating in the study will be stored on a secure server in a password-protected account and will only be available to the researchers.

Will you receive anything for being in this study?

Yes. You will be given \$30 for your participation. Early termination will result in \$20 compensation.

Will it cost you anything to be in this study?

It will cost you nothing to be in this study, other than time to come to the UNC campus location and participate in the study.

What if you are a UNC student?

You may choose not to be in the study or to stop being in the study before it is over

at any time. This will not affect your class standing at UNC-Chapel Hill. You will not be offered or receive any special consideration if you take part in this research.

What if you have questions about this study?

You have the right to ask, and have answered, any questions you may have about this research. If you have questions, or concerns, you should contact the researchers listed on the first page of this form.

What if you have questions about your rights as a research participant?

All research on human volunteers is reviewed by a committee that works to protect your rights and welfare. If you have questions or concerns about your rights as a research subject you may contact, anonymously if you wish, the Institutional Review Board at 919-966-3113 or by email to IRB subjects@unc.edu.

Appendix I – Consent Form for Study 2

Consent Form							
Title of Study: Modeling the perception of online coll IRB Study Number: 12-0255 Consent Form Version Date: March 1st, 2012 Principal Investigator: Felix Portnoy	ntent						
Participant's Agreement:							
I have read the information provided to me in the information provided to							
Signature of Research Participant	Date						
Printed Name of Research Participant							
Signature of Person Obtaining Consent	 Date						
Felix Portnoy Printed Name of Person Obtaining Consent							

REFERENCES

- Appel, V. (1971). On Advertising Wear Out. *Journal of Advertising Research*, 11(1), 11–13.
- Ashkan, A., Clarke, C., Agichtein, E., & Guo, Q. (2009). Classifying and Characterizing Query Intent. In M. Boughanem, C. Berrut, J. Mothe, & C. Soule-Dupuy (Eds.), *Advances in Information Retrieval*, Lecture Notes in Computer Science (Vol. 5478, pp. 578–586). Springer Berlin / Heidelberg. Retrieved from http://www.springerlink.com.libproxy.lib.unc.edu/content/6j75165rw8386 055/abstract/
- Barlow, T., & Wogalter, M. S. (1993). Alcoholic Beverage Warnings in Magazine and Television Advertisements. *Journal of Consumer Research*, 20(1), 147–156.
- Bayles, M. E. (2002). Designing online banner advertisements: should we animate? *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves* (pp. 363-366). Minneapolis, Minnesota, USA: ACM.
- Benway, J. P. (1998). Banner Blindness: The Irony of Attention Grabbing on the World Wide Web. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 42, 463–467.
- Berdie, D. R. (1992). The Yellow Pages Report: A comprehensive guide for advertisers. *St. Paul, MN: Consumer Review Systems*.
- Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Perception & Psychophysics*, 8(5-A), 279–286.
- Berntson, G. G., Quigley, K. S., & Lozano, D. (2007). Cardiovascular psychophysiology. In J. T Cacioppo & L. G. Tassinary (Eds.), *Handbook of psychophysiology* (pp. 182–210). Cambridge University Press.
- Bolt, R. A. (1982). Eyes at the interface. *Proceedings of the 1982 conference on Human factors in computing systems* (p. 360-362).
- Boucsein, W. (1992). *Electrodermal activity*. Plenum Pub Corp.
- Brahn, M. (1901). Experimented Beitrage zur Gefuhlslehre. Phil. Stud, 18, 127-187.
- Briggs, R., & Hollis, N. (1997). Advertising on the Web: Is there response before click-through? *Journal of Advertising Research*, *37*(2), 33–45.

- Broadbent, D. E. (1952). Listening to one of two synchronous messages. *Journal of Experimental Psychology*, 44(1), 51-55.
- Broadbent, D. E. (1954). The role of auditory localization in attention and memory span. *Journal of Experimental Psychology*, 47(3), 191-196.
- Broadbent, D. E. (1958). *Perception and communication*. Elmsford, NY, US: Pergamon Press, Inc.
- Broadbent, Donald Eric. (1971). Decision and stress. Academic Press.
- Broadbent, Donald E. (1982). Task combination and selective intake of information. *Acta Psychologica*, *50*(3), 253-290.
- Broder, A. (2002). A taxonomy of web search. *SIGIR Forum*, *36*(2), 3–10. doi:10.1145/792550.792552
- Buchanan, B., & Morrison, D. G. (1988). A stochastic model of list falloff with implications for repeat mailings. *Journal of Direct Marketing*, 2(3), 7-15.
- Burke, M. C., & Edell, J. A. (1986). Ad Reactions over Time: Capturing Changes in the Real World. *The Journal of Consumer Research*, *13*(1), 114–118.
- Burke, M., Hornof, A., Nilsen, E., & Gorman, N. (2005). High-cost banner blindness: Ads increase perceived workload, hinder visual search, and are forgotten. *ACM Trans. Comput.-Hum. Interact.*, *12*(4), 423–445.
- Calisir, F., & Karaali, D. (2008). The impacts of banner location, banner content and navigation style on banner recognition. *Comput. Hum. Behav.*, *24*(2), 535–543. doi:http://dx.doi.org.libproxy.lib.unc.edu/10.1016/j.chb.2007.02.019
- Card, S. K. (1984). Visual search of computer command menus. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 97–108).
- Cassidy, S., & Eachus, P. (2002). Developing the computer user self-efficacy (CUSE) scale: Investigating the relationship between computer self-efficacy, gender and experience with computers. *Journal of Educational Computing Research*, 26(2), 133–153.
- Chatterjee, P., Hoffman, D. L., & Novak, T. P. (2003). Modeling the Clickstream: Implications for Web-Based Advertising Efforts. *Marketing Science*, 22(4), 520–541.
- Chen, M. C., Anderson, J. R., & Sohn, M. H. (2001). What can a mouse cursor tell us more?: correlation of eye/mouse movements on web browsing. *CHI '01*

- extended abstracts on Human factors in computing systems, CHI EA '01 (pp. 281–282). New York, NY, USA: ACM. doi:10.1145/634067.634234
- Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America*, *25*(5), 975-979.
- Cho, C.-H., & Cheon, H. J. (2004). Why Do People Avoid Advertising on the Internet? *Journal of Advertising*, *33*(4), 89–97.
- Clark, A. (2009, August 6). Rupert Murdoch plans charge for all news websites by next summer. *The Guardian*. Retrieved November 7, 2009, from http://www.guardian.co.uk/media/2009/aug/06/rupert-murdoch-website-charges
- Cowen, L., Ball, L. J., & Delin, J. (2002). An eye movement analysis of webpage usability. *People and Computers XVI: Memorable Yet Invisible, Proceedings of HCI* (pp. 317–335).
- D'Angelo, F. (2009, October 26). Happy Birthday, Digital Advertising! *Advertising Age*. Retrieved November 7, 2009, from http://adage.com/digitalnext/article?article_id=139964
- Danaher, P. J., & Mullarkey, G. W. (2003). Factors affecting online advertising recall: A study of students. *Journal of Advertising Research*, 43(03), 252–267.
- Danaher, P. J., Mullarkey, G. W., & Essegaier, S. (2006). Factors Affecting Web Site Visit Duration: A Cross-Domain Analysis. *Journal of Marketing Research*, 43(2), 182–194.
- Darrow, C. W. (1929). Differences in the physiological reactions to sensory and ideational stimuli. *Psychological Bulletin*, *26*(4), 185-201.
- Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological review*, *70*(1), 80-90.
- Diao, F., & Sundar, S. Shyam. (2004). Orienting Response and Memory for Web Advertisements:: Exploring Effects of Pop-Up Window and Animation. *Communication Research*, *31*(5), 537-567.
- Drèze, X., & Hussherr, F.-X. (2003). Internet advertising: Is anybody watching? *Journal of Interactive Marketing*, 17(4), 8-23.
- Eachus, P., & Cassidy, S. (2006). Development of the web users self-efficacy scale (WUSE). *Issues in Informing Science and Information Technology Journal*, *3*, 199–209.

- Eachus, P., Hogg, P., & Cassidy, S. (2006). Further development of the web user self-efficacy scale (WUSE).
- Eveland, W. P., & Dunwoody, S. (2002). An investigation of elaboration and selective scanning as mediators of learning from the Web versus print. *Journal of Broadcasting & Electronic Media*, 46(1), 34–53. doi:10.1207/s15506878jobem4601_3
- Field, A. (2009). *Discovering Statistics Using SPSS*. SAGE Publications.
- Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47, 381–391.
- Fitts, P., M., Jones, R., E., & Milton, J., L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. In N. Moray (Ed.), *Ergonomics: Psychological mechanisms and models in ergonomics* (pp. 56–66). Ergonomics: Psychological mechanisms and models in ergonomics: Taylor & Francis US.
- Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, 18(4), 1030-1044.
- Fry, T. R. L. (1996). Advertising wearout in the Transport Accident Commission road safety campaigns. *Accident Analysis & Prevention*, *28*(1), 123-129.
- Go, G. (2008). Best Performing Banner Ad Sizes. *About.com*. Retrieved September 26, 2011, from http://onlinebusiness.about.com/b/2008/09/17/best-performing-banner-ad-sizes.htm
- Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics*, 24(6), 631–645.
- Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N., & Wichansky, A. M. (2002). Eye tracking in web search tasks: design implications. *Proceedings of the 2002 symposium on Eye tracking research & applications* (pp. 51–58).
- Goldeberg, H., & Wichansky, A. M. (2003). Eye tracking in usability evaluation: A practitioner's guide. In J. Hyönä & R. Radach (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 493-516).
- Heinonen, K., & Strandvik, T. (2007). Consumer responsiveness to mobile marketing. *International Journal of Mobile Communications*, *5*(6), 603–617.

- Hillstrom, A. P., & Yantis, S. (1994). Visual motion and attentional capture. *Attention, Perception, & Psychophysics*, *55*(4), 399–411.
- Horn, G., Hinde, R. A., & Centre, K. C. (University of C. R. (1970). *Short-term changes in neural activity and behaviour*. CUP Archive.
- Hutchinson, T., White, K., Martin, W., Reichert, K., & Frey, L. (1989). Human-computer interaction using eye-gaze input. *IEEE Transactions on Systems, Man, and Cybernetics, 19*(6), 1527–1534.
- Hyönä, J., Tommola, J., & Alaja, A. M. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *The Quarterly Journal of Experimental Psychology Section A*, 48(3), 598–612.
- IAB. (2011). IAB Internet Advertising Revenue Report.
- Jacob, R., & Karn, K. (2003). Commentary on Section 4. Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: cognitive and applied aspects of eye movement research* (pp. 573–606). Elsevier.
- James, W. (1890). The principles of psychology. H. Holt and company.
- Janiszewski, C. (1998). The Influence of Display Characteristics on Visual Exploratory Search Behavior. *The Journal of Consumer Research*, 25(3), 290–301.
- Jansen, B. J., Booth, D. L., & Spink, A. (2008). Determining the informational, navigational, and transactional intent of Web queries. *Inf. Process. Manage.*, 44(3), 1251–1266. doi:10.1016/j.ipm.2007.07.015
- Jenkin, M. R. M., & Harris, L. R. (Eds.). (2001). Vision and Attention. Springer.
- Jones, M. Y., Pentecost, R., & Requena, G. (2005). Memory for advertising and information content: Comparing the printed page to the computer screen. *Psychology & Marketing*, 22(8), 623–648. doi:10.1002/mar.20077
- Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Attention, Perception, & Psychophysics*, *43*(4), 346–354.
- Jung, C. G. (1906). Psychoanalysis and association experiments. *Collected Works*, 2.
- Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*(4), 441–480.
- Kaiser, S. E. (2007). Banner Ads: The Best Performing Banner Ad Sizes Ad Formats Locations and Colors. *WebsiteTips.com*. Retrieved September 26, 2011, from http://websitetips.com/articles/marketing/banneradsizes/

- Kalsbeek, J. W. H., & Ettema, J. H. (1963). Scored regularity of the heart rate pattern and the measurement of perceptual or mental load. *Ergonomics*, 6(3), 306–307.
- Kalyanaraman, S. (in press). Cogito interruptus: The role of interruption type, interruption content, and interruption frequency in online information processing. *Journal of the American Society for Information Science and Technology*.
- Karat, C.-M., Halverson, C., Horn, D., & Karat, J. (1999). Patterns of entry and correction in large vocabulary continuous speech recognition systems. *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*, CHI '99 (p. 568–575). New York, NY, USA.
- Kramer, A., F. (1991). Physiological metrics of menral workload: A review of recent progress. In D. L. Damos (Ed.), *Multiple-task performance* (pp. 279–321). CRC Press.
- Kramer, A., F., & Parasuraman, R. (2007). Neroergonomics: Application of Neuroscience to Human Factors. In John T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology*. Cambridge University Press.
- Kuhmann, W., Boucsein, W., Schaefer, F., & Alexander, J. (1987). Experimental investigation of psychophysiological stress-reactions induced by different system response times in human-computer interaction*. *Ergonomics*, *30*(6), 933–943.
- Lacey, J. I., Kagan, J., Lacey, B. C., & Moss, H. A. (1963). The visceral level: Situational determinants and behavioral correlates of autonomic response patterns. In P. Knapp (Ed.), *Expression of the emotions in man* (pp. 161–196). New York: International Universities Press.
- Lai, J., & Vergo, J. (1997). MedSpeak: report creation with continuous speech recognition. *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '97 (p. 431–438). New York, NY, USA.
- Landis, C., & Hunt, W. A. (1935). The conscious correlates of the galvanic skin response. *Journal of Experimental Psychology*, *18*(5), 505–529.
- Lang, A., Borse, J., Wise, K., & David, P. (2002). Captured by the World Wide Web: Orienting to Structural and Content Features of Computer-Presented Information. *Communication Research*, 29(3), 215-245
- Lang, A., Potter, R. F., & Bolls, P. (2008). Where Psychophsiology Meets the Media. In J. Bryant & M. B. Oliver (Eds.), *Media Effects: Advances in Theory and Research* (pp. 185–206). Taylor & Francis.

- Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology*, *30*, 261–261.
- Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology-General*, 133(3), 339–353.
- Leigh, J. H. (1991). Information Processing Differences among Broadcast Media: Review and Suggestions for Research. *Journal of Advertising*, 20(2), 71–75.
- Lewandowski, D., Drechsler, J., & von Mach, S. (in press). Deriving Query Intents From Web Search Engine Queries. *Journal of the American Society for Information Science & Technology*.
- Lockerd, A., & Mueller, F. M. (2002). LAFCam: Leveraging affective feedback camcorder. *CHI '02 extended abstracts on Human factors in computing systems* (pp. 574–575). Minneapolis, Minnesota, USA: ACM. Retrieved from http://portal.acm.org.libproxy.lib.unc.edu/citation.cfm?id=506490&coll=GUIDE&dl=GUIDE&CFID=90829335&CFTOKEN=17059806&ret=1#Fulltext
- Loftus, E. F. (1974). On reading the fine print. *The Quarterly Journal of Experimental Psychology*, *26*(2), 324-324.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological review*, *95*(4), 492.
- Lohse, G. L. (1997). Consumer eye movement patterns on yellow pages advertising. *Journal of Advertising*, *26*(1), 61–73.
- Lutz, K. A., & Lutz, R. J. (1977). Effects of interactive imagery on learning: Application to advertising. *Journal of Applied Psychology*, *62*(4), 493–498.
- MacKay, D. M. (1973). Visual stability and voluntary eye movements. In R. Jung (Ed.), *Handbook of sensory physiology* (Vol. 8/3, pp. 307–331). Berlin: Springer-Verlag.
- MacKenzie, S. B., Lutz, R. J., & Belch, G. E. (1986). The Role of Attitude Toward the Ad as a Mediator of Advertising Effectiveness: A Test of Competing Explanations. *Journal of Marketing Research (JMR)*, 23(2), 130–143.
- MacLeod, C. M. (1992). The Stroop task: The gold standard of attentional measures. *Journal of Experimental Psychology: General*, 121(1), 12.
- Mandryk, R. L., Inkpen, K. M., & Calvert, T. W. (2006). Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour & Information Technology*, 25(2), 141.

- Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communications of the ACM*, 49(4), 46.
- Maxwell, S. E., & Delaney, H. D. (2004). *Designing experiments and analyzing data: a model comparison perspective* (2nd ed.). Psychology Press.
- Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, *11*(1), 56.
- Moray, N. (1969). *Attention: selective processes in vision and hearing*. Academic Press.
- Müller, J., Wilmsmann, D., Exeler, J., Buzeck, M., Schmidt, A., Jay, T., & Krüger, A. (2009). Display Blindness: The Effect of Expectations on Attention towards Digital Signage. *Pervasive Computing*, 1–8.
- Mundy-Castle, A. C., & McKiever, B. L. (1953). The psychophysiological significance of the galvanic skin response. *Journal of Experimental Psychology*, 46(1), 15-24.
- Nakayama, M., Takahashi, K., & Shimizu, Y. (2002). The act of task difficulty and eyemovement frequency for the Oculo-motor indices'. *Proceedings of the 2002 symposium on Eye tracking research & applications* (p. 42).
- Neumann, E., & Blanton, R. (1970). The early history of electrodermal research. *Psychophysiology*, *6*(4), 453–475.
- Newstead, S. E., & Dennis, I. (1979). Lexical and grammatical processing of unshadowed messages: A re-examination of the Mackay effect. *Quarterly Journal of Experimental Psychology*, 31(3), 477.
- Nielsen, J. (1997). How Users Read on the Web. *useit.com*. Retrieved April 29, 2012, from http://www.useit.com/alertbox/9710a.html
- Noton, D., & Stark, L. (1971). Eye movements and visual perception. *Scientific American*, 224(6), 35–43.
- Ottati, W. L., Hickox, J. C., & Richter, J. (1999). Eye scan patterns of experienced and novice pilots during visual flight rules (VFR) navigation. *Human Factors and Ergonomics Society Annual Meeting Proceedings* (Vol. 43, pp. 66–70).
- Oviatt, S., Coulston, R., & Lunsford, R. (2004). When do we interact multimodally?: cognitive load and multimodal communication patterns. *Proceedings of the 6th international conference on Multimodal interfaces* (pp. 129–136).
- Owens, J. W., Chaparro, B. S., & Palmer, E. M. (2011). Text Advertising Blindness: The New Banner Blindness? *Journal of Usability Studies*, 6(3), 172–197.

- Pagendarm, M., & Schaumburg, H. (2001). Why Are Users Banner-Blind? The Impact of Navigation Style on the Perception of Web Banners. *Journal of Digital Information*, *2*(1). Retrieved from http://journals.tdl.org/jodi/article/viewArticle/36/38
- Pan, B., Hembrooke, H. A., Gay, G. K., Granka, L. A., Feusner, M. K., & Newman, J. K. (2004). The determinants of web page viewing behavior: an eye-tracking study. *Proceedings of the 2004 symposium on Eye tracking research* \& applications (pp. 147–154). San Antonio, Texas: ACM. Retrieved from http://portal.acm.org.libproxy.lib.unc.edu/citation.cfm?id=968391
- Pashler, H. E. (1999). Attention and Perception. *The Psychology of Attention* (pp. 35–262). MIT Press.
- Pashler, H. E. (2001). Involuntary orienting to flashing distractors in delayed search? In C. L. Folk & B. S. Gibson (Eds.), *Attraction, distraction and action: multiple perspectives on attentional capture* (Vol. 133, pp. 77–92). Elsevier.
- Pechmann, C., & Stewart, D. W. (1988). Advertising Repetition: A Critical Review of Wearin and Wearout. *Current Issues & Research in Advertising*, 11(2), 285.
- Peters, J. W. (2011, March 17). New York Times to Impose Fees for Web Readers on March 28. *The New York Times*. Retrieved from http://www.nytimes.com/2011/03/18/business/media/18times.html?_r=1 &src=me&ref=general
- Picard, R. W. (2000). Affective Signals and Systems. *Affective Computing* (pp. 141–164). MIT Press.
- Poole, A., & Ball, L. J. (2000). Eye tracking in human-computer interaction and usability research: current status and future prospects. In C. Ghaoui (Ed.), *Encyclopedia of human computer interaction* (pp. 211–219). Idea Group.
- Posner, M. I., & Snyder, C. R. R. (1975). Facilitation and inhibition in the processing of signals. *Attention and performance V*, 669–682.
- Prideaux, E. (1920). The psychogalvanic reflex: A review. *Brain*, 43(1), 50–73.
- Quinlan, P. T. (2003). Visual feature integration theory: Past, present, and future. *Psychological Bulletin*, *129*(5), 643.
- Rau, P. P.-L., & Chen, D. (2006). Effects of watermark and music on mobile message advertisements. *International Journal of Human-Computer Studies*, 64(9), 905-914.
- Rayner, Keith. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372-422.

- Rethans, A. J., Swasy, J. L., & Marks, L. J. (1986). Effects of television commercial repetition, receiver knowledge, and commercial length: a test of the two-factor model. *Journal of Marketing Research*, 23(1), 50–61.
- Reynolds, D. (1964). Effects of double stimulation: Temporary inhibition of response. *Psychological Bulletin*, *62*(5), 333-347.
- Richardson, D., & Spivey, M. (2004).In G. E. Wnek & G. L. Bowlin (Eds.), *Encyclopedia of Biomaterials and Biomedical Engineering: L-Z* (pp. 568–572). Informa Health Care.
- Rodden, K., Fu, X., Aula, A., & Spiro, I. (2008). Eye-mouse coordination patterns on web search results pages. *CHI '08 extended abstracts on Human factors in computing systems*, CHI EA '08 (pp. 2997–3002). New York, NY, USA: ACM. doi:10.1145/1358628.1358797
- Rowe, D. W., Sibert, J., & Irwin, D. (1998). Heart rate variability: indicator of user state as an aid to human-computer interaction. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 480–487).
- Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eyetracking protocols. *Proceedings of the 2000 symposium on Eye tracking* research & applications (pp. 71–78).
- Sawyer, A. (1981). Repetition, Cognitive Responses and Persuasion. In T. M. Ostrom, R. E. Petty, & T. C. Brock (Eds.), *Cognitive Responses in Persuasion*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Scheirer, J., Fernandez, R., Klein, J., & Picard, R. W. (2002). Frustrating the user on purpose: a step toward building an affective computer. *Interacting with Computers*, *14*(2), 93-118.
- Scherer, K. R. (2001). Appraisal considered as a process of multi-level sequential checking. In K. R. Scherer & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 92–120). New York: Oxford University Press, USA.
- Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. University of Illinois Press.
- Shapiro, S., MacInnis, D. J., & Heckler, S. E. (1997). The effects of incidental ad exposure on the formation of consideration sets. *Journal of Consumer Research*, 94–104.
- Simon, H. A. (1967). MOTIVATIONAL AND EMOTIONAL CONTROLS OF COGNITION. *Psychological Review*, 74(1), 29-39.

- Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *PERCEPTION-LONDON-*, *28*, 1059–1074.
- Sirevaag, E. J., Kramer, A. F., Wickens, C. D, Reisweber, M., Strayer, D. L., & Grenell, J. F. (1993). Assessment of pilot performance and mental workload in rotary wing aircraft. *Ergonomics*, 36(9), 1121-1140.
- Smith, R. E., & Buchholz, L. M. (1991). Multiple Resource Theory and Consumer Processing of Broadcast Advertisements: An Involvement Perspective. *Journal of Advertising*, 20(3), 1–7.
- Sokolov, E. N. (1963). Perception and the conditioned reflex. Pergamon Press.
- Spool, J. M., Scanlon, T., Schroeder, W., Snyder, C., & DeAngelo, T. (1997). Web site usability: a designer's guide. Morgan Kaufmann.
- Steinberg, J. (2011, April 23). Click Through Rates and The New 90-10-1 Rule for Ads | Jon Steinberg. *jonsteinberg.com*. Retrieved May 2, 2011, from http://jonsteinberg.com/2011/04/23/click-rates/
- Stern, R. M., Ray, W. J., & Quigley, K. S. (2001). Electrodermal Activity. *Psychophysiological recording* (pp. 206–220). Oxford University Press, USA.
- Stevens, H. C. (1905). A plethysmographic study of attention. *The American Journal of Psychology*, *16*(4), 409–483.
- Stroop, J. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662.
- Styles, E. A. (2006). *The psychology of attention*. Psychology Press.
- Sundar, S. S, & Kalyanaraman, S. (2004). Arousal, memory, and impression-formation effects of animation speed in web advertising. *Journal of Advertising*, 33(1), 7–17.
- Sundar, S. S, & Wagner, C. B. (2002). The world wide wait: Exploring physiological and behavioral effects of download speed. *Media Psychology*, 4(2), 173–206.
- Sundar, S. Shyam, Narayan, S., Obregon, R., & Uppal, C. (1998). Does Web Advertising Work? Memory for Print Vs. Online Media. *Journalism & Mass Communication Quarterly*, 75(4), 822–835. doi:10.1177/107769909807500414
- Tattersall, A. J., & Hockey, G. R. (1995). Level of Operator Control and Changes in Heart Rate Variability during Simulated Flight Maintenance. *Human Factors*, *37*(4) 682-698.
- Theron, P. A. (1948). Peripheral vasomotor reactions as indices of basic emotional tension and lability. *Psychosomatic Medicine*, *10*(6), 335-346.

- Thomas, L. C., & Wickens, C. D. (2004). Eye-tracking and individual differences in offnormal event detection when flying with a synthetic vision system display. *Human Factors and Ergonomics Society Annual Meeting Proceedings* (Vol. 48, pp. 223–227).
- Thorley, P., Hellier, E., & Edworthy, J. (2001). Habituation effects in visual warnings. In M. Hanson (Ed.), *Contemporary ergonomics 2001* (pp. 223–228). CRC Press.
- Thorley, P., Hellier, E., Edworthy, J., & Stephenson, D. (2002). Orienting response reinstatement in text text and pictorial warnings. In P. T. McCabe (Ed.), *Contemporary Ergonomics 2002* (pp. 447–451). CRC Press.
- Thüring, M., & Mahlke, S. (2007). Usability, aesthetics and emotions in human–technology interaction. *International Journal of Psychology*, 42(4), 253-264.
- Titchener, E. B. (1908). *Lectures on the elementary psychology of feeling and attention*. Adamant Media Corporation.
- Tong, H., & Fisher, R. (1984). *Progress Report on an Eye-Slaved Area-of-Interest Visual Display* (No. AFHRL-TR-84-36). Brooks Air Force Base, Texas1: Air Force Human Resources Laboratory.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, *95*(1), 15-48.
- Treisman, A. M. (1964a). Verbal cues, language, and meaning in selective attention. *The American Journal of Psychology*, *77*(2), 206–219.
- Treisman, A. M. (1964b). The effect of irrelevant material on the efficiency of selective listening. *The American Journal of Psychology*, 77(4), 533–546.
- Treisman, A. M. (1964c). Selective attention in man. *British Medical Bulletin, 20*(1), 12-16.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97-136.
- Treisman, Anne, & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*(1), 107–141. doi:10.1016/0010-0285(82)90006-8
- Tsang, M. M., Ho, S. C., & Liang, T. P. (2004). Consumer attitudes toward mobile advertising: An empirical study. *International Journal of Electronic Commerce*, 8(3), 65–78.

- Tullis, T., Siegel, M., & Sun, E. (2009). Are people drawn to faces on webpages? *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, CHI EA '09 (p. 4207–4212). New York, NY, USA.
- Valiente, R. (1973). Mechanical correlates of ad recognition. *Journal of Advertising Research*, 13(3), 13–18.
- Van Der Merwe, A. B., & Theron, P. A. (1947). A new method of measuring emotional stability. *Journal of General Psychology*, *37*, 109–23.
- Van Orden, K. F., Divita, J., & Shim, M. J. (1993). Redundant use of luminance and flashing with shape and color as highlighting codes in symbolic displays. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 35(2), 195–204.
- van Ravenswaaij-Arts, C. M. A., Kollee, L. A. A., Hopman, J. C. W., Stoelinga, G. B. A., & van Geijn, H. P. (1993). Heart Rate Variability. *Annals of Internal Medicine*, 118(6), 436-447.
- Vitense, H. S., Jacko, J. A., & Emery, V. K. (2002). Multimodal feedback: establishing a performance baseline for improved access by individuals with visual impairments. *Proceedings of the fifth international ACM conference on Assistive Technologies* (pp. 49–56).
- Ward, R. D., & Marsden, P. H. (2003). Physiological responses to different WEB page designs. *International Journal of Human-Computer Studies*, *59*(1-2), 199–212.
- Ward, R. D., Marsden, P. H., Cahill, B., & Johnson, C. (2002). Physiological responses to well-designed and poorly designed interfaces. *Proceedings of CHI 2002 Workshop on Physiological Computing*, Minneapolis, MN.
- Weber, E. (1846). Wagner's Handwörterbuch der Physiologie. Braunschweig: Vieweg.
- Wedel, M., & Pieters, R. (2007). A review of eye-tracking research in marketing. *Review of marketing research*, *4*, 123–47.
- Wen, Z., Zhu, Y., & Peng, Z. (2009). Survey on Web Image Content-Based Filtering Technology. *Information Science and Engineering (ICISE), 2009 1st International Conference on* (pp. 1463 –1466). doi:10.1109/ICISE.2009.1152
- Werbin, F., & Roska, B. (2007, April). The Movies in Our Eyes. *Scientific American*, 72–79.
- Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63–101). New York: Academic Press.

- Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177.
- Wickens, Christopher D., & McCarley, J. S. (2008). *Applied attention theory*. CRC Press.
- Wilson, G. M., & Sasse, M. A. (2000). Do users always know what's good for them? Utilising physiological responses to assess media quality. *People and Computers XIV–Usability or Else! Proceedings of HCI 2000* (pp. 327–341).
- Wise, K., & Reeves, B. (2007). The Effect of User Control on the Cognitive and Emotional Processing of Pictures. *Media Psychology*, *9*(3), 549-566.
- Wogalter, M. S, Magurno, A. B., Carter, A. W., Swindell, J. A., Vigilante, W. J., & Daurity, J. G. (1995). Hazard associations of warning header components. *Human Factors and Ergonomics Society Annual Meeting Proceedings* (Vol. 39, pp. 979–983).
- Wogalter, Michael S., & Brelsford, J. W. (1994). Incidental Exposure to Rotating Warnings on Alcoholic Beverage Labels. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 38, 374–378.
- Wogalter, Michael S., DeJoy, D. M., & Laughery, K. R. (1999). *Warnings and risk communication*. CRC Press.
- Wolfe, J. M. (1994). Guided search 2. 0. A revised model of visual search. *Psychonomic Bulletin & Review*, *1*(2), 202–238.
- Wroblewski, L. (2008). Web form design. Rosenfeld Media.
- Yesilada, Y., Jay, C., Stevens, R., & Harper, S. (2008). Validating the use and role of visual elements of web pages in navigation with an eye-tracking study. *Proceeding of the 17th international conference on World Wide Web*, WWW '08 (pp. 11–20). Beijing, China: ACM.
- Zeff, R. L., & Aronson, B. (1999). *Advertising on the Internet* (2nd ed.). New York: Wiley,.
- Zoneff, P., & Meumann, E. (1902). Uber Begleiterscheinungen psychischer Vorgange bei Athem und Pulse. *Phil Stud.*, *18*, 1–113.