GETTING THE GIST OF WEBSITES: EXPLORING THE EFFECTS OF DISPLAY DURATION, SIZE, AND RESOLUTION

A Dissertation by

Justin Wade Owens

Master of Arts, Wichita State University, 2010

Bachelor of Business Administration, Wichita State University, 2001

Submitted to the Department of Psychology
and the faculty of the Graduate School of
Wichita State University
in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

May 2013

GETTING THE GIST OF WEBSITES: EXPLORING THE EFFECTS OF DISPLAY DURATION, SIZE, AND RESOLUTION

The following faculty members have examined the final copy of this dissertation for form and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Doctor of Philosophy, with a major in Psychology.

Barbara Chaparro, Committee Chair

Evan Palmer, Committee Member

Rui Ni, Committee Member

Alex Chaparro, Committee Member

Jim Wolff, Committee Member

Accepted for the Fairmount College of Liberal Arts and Sciences

Ron Matson, Interim Dean

Accepted for the Graduate School

Abu S. M. Masud, Interim Dean

ABSTRACT

Users can make judgments about web pages in a glance. Outside of a few studies, no known research has examined what semantic information can be extracted from a web page within a single fixation, but the scene perception literature provides a possible framework for understanding how viewers can extract diverse semantic information from scenes in a glance. The purpose of this dissertation was to explore whether semantic information about a web page could be extracted within a single fixation and to explore the effects of size and resolution on extracting this information. Initially, the classification of web pages was explored, which provided web page stimuli for the following two studies. Using a rapid serial visual presentation (RSVP) paradigm, the first study explored whether certain semantic categories of websites (i.e., news, search, shopping, and social networks/blogs) could be detected from stream of web page stimuli that were presented briefly. Natural scenes, which have been shown to be detectable within a single fixation in the literature, served as a baseline for comparison. The second study examined the effects of stimulus size and resolution on observers' ability to detect the presence of a certain website category using similar methods.

Results from this research showed that users have conceptual models of websites. These models allowed detection of web pages from a fixation's worth of stimulus exposure when provided with additional time for processing information. For the website categories other than search, detection performance decreased significantly when web elements were no longer discernible due to decreases in size and/or resolution. The implications of this research are that website conceptual models rely heavily on page elements and less on the spatial relationship between these elements.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

Table Page

LIST OF TABLES

LIST OF TABLES (continued)

Table Page

LIST OF TABLES (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

Figure Page

CHAPTER 1

INTRODUCTION

The existence of website layout conventions, or where users expect web objects to be located, has been demonstrated extensively (Bernard, 2001; Bernard, 2003; Bernard & Sheshadri, 2004; Di Nocera, Capponi, and Ferlazzo, 2004; Granka, Hembrooke, & Gay, 2006; Owens, Chaparro, & Palmer, 2011; Owens, Palmer, & Chaparro, in press; Roth, Schmutz, Pauwels, Bargas-Avila, & Opwis, 2010; Shaikh, Chaparro, & Joshi, 2006; Shaikh & Lenz, 2006). For instance, users expect navigation on the left side or top of a website, while advertising can be expected on the top or right side (Bernard, 2001; Bernard & Sheshadri, 2004; Shaikh, et al., 2006; Shaikh & Lenz, 2006). Such findings have been shown to be cross-cultural (Bernard & Sheshadri, 2004; Shaikh, et al., 2006), exist for specific types of websites (Roth, et al., 2010), and affected by users' experience and expertise (Di Nocera, et al., 2004; Roth, et al., 2010). While website layout conventions do exist, research has shown also that users adapt within a few page views to violations of these conventions, but sometimes at costs in terms of usability (McCarthy, Sasse, and Riegelsberger, 2003; Owens, et al., in press; Santa-Maria, L. & Dyson, 2008; Tzanidou, Petre, Minocha, & Grayson, 2005).

Few studies have approached examining what information can be derived from web pages from a mere glance. Previously exploration into what can be perceived from a web page within a few hundred milliseconds has been limited to visual appeal, trustworthiness, and perceived usability (Albert, Gribbons, & Almadas, 2009; Lindgaard, Fernandes, Dudek, & Brown, 2006; Lindgaard, Dudek, Sen, Sumegi, & Noonan, 2011; Thielsch & Hirschfeld, 2012; Tuch, Presslaber, Stocklin, Opwis, & Bargas-Aliva, 2012). The visual appeal of web pages can be

judged in 50 ms almost as consistently in 500 ms (Lindgaard, et al., 2006; Lindgaard, et al., 2011). Such findings regarding visual appeal have also been consistent with unlimited viewing (Lindgaard, et al., 2006). With exposures as low as 17 ms, visual appeal has been shown to be correlated with the prototypicality and visual complexity of a web page (Tuch, et al., 2012). Moreover, visual appeal of websites has been shown to be related to low spatial frequencies of web pages when they were displayed for only 50 ms (Thielsch, & Hirschfeld, 2012). With similar display durations (i.e., 50 ms), both trust and perceived usability could be rated (Albert, et al., 2009; Lindgaard, et al., 2011). At 50 ms, these exposures were substantially less than the duration of a fixation (Rayner, 2009).

The findings of these studies raise the question that if these types of judgments can reliably be made in a glance and users have well defined conventions for websites, how are websites perceived and represented cognitively? Currently, there is no known research exploring the perceptual and cognitive representations of web pages in these short durations and how quickly they can be accessed outside of these few studies. However, this question may be explored by using the scene perception literature to provide a well-established framework.

A common theme across the exploration of scene perception has been how visual scenes can be classified. Scenes have been classified into superordinate, basic categories, and subordinate categories (Rosch, 1978; Tversky & Hemenway, 1983). An example with increasing abstraction would be sandy versus rocky beaches (i.e., subordinate), which both are beaches (i.e., basic-level), which in turn are all outdoor scenes (i.e., superordinate). Essentially, scenes within basic categories are most similar, while exhibiting the most dissimilarity with other basic categories. Superordinate classifications are more abstract and share fewer features between

classifications. In essence, basic level classifications contain more detail than superordinate classification.

In several scene perception studies, this classification method has been applied extensively to scenes. Typically, the scenes can be classified into superordinate and basic level classifications, which are related through global properties and that superordinate classifications can be distinguished prior to basic-level classifications (Fei-Fei, Iyer, Koch, & Perona, 2007; Greene & Oliva, 2009a; Loschky & Larson, 2010; Oliva & Torralba, 2001). For example, a forest is a basic-level classification while the superordinate classification could be either outdoor or natural scenes. In such examples, research has shown that the superordinate categories can be processed with higher performance and distinguished earlier than basic-level categories (Fei-Fei, et al., 2007; Loschky & Larson, 2010). Some differences have been found between superordinate categories where outdoor scenes can be processed prior to indoor scenes (Fei-Fei, et al., 2007).

Such findings support global-to-local processing of scenes, where the global information of a scene is processed first, after which local information is processed (Navon, 1977). This is analogous to processing the forest prior to the trees. Oliva & Torralba (2001) proposed one such model called the spatial envelope, where global properties of a scene can classify scenes into superordinate as well as basic categories. These properties included naturalness, roughness, openness, ruggedness, and expansion of the scenes. Additionally, changes to global statistics have been shown to affect the perception of scenes (Joubert, Rousselet, Fabre-Thorpe, & Fize, 2009). Some research has suggested that superordinate categories have less bias than basic categorizations (Loschky & Larson, 2010) and when scenes share global

properties or do not have a distinct global property, correctly distinguishing between scenes becomes more difficult (Greene & Oliva, 2009b; Loschky & Larson, 2010).

Several scene perception theories incorporate processing of global information for the recognition of objects within a scene. These include the perceptual schema model, the priming model, and contextual guidance model (Friedman, 1979; Henderson & Hollingsworth, 1999; Oliva & Torralba, 2007; Torralba, Oliva, Castelhano, & Henderson, 2006). Such concepts also have been integrated into other theories related to vision, such as visual search models (Wolfe, Võ, Evans, & Greene, 2011).

Within the scene perception literature, a theory called scene gist has been researched extensively. Scene gist has been described as the rapidly extracted meaning of a scene that occurs within a fixation, possibly either in the absence of or with little attention, and has shown to be related to the global processing of information (Fei-Fei, vanRullen, Koch, & Perona, 2002; Fei-Fei, et al., 2007; Greene and Oliva, 2009a; Greene and Oliva, 2009b; Larson & Loschky, 2009; Oliva, 2005; Potter, 1975; Potter, 1976; Intraub, 1980; Intraub, 1981). The information contained within scene gist has been thought to consist of a semantic label, a limited number of objects, and the spatial layout of objects (Oliva & Torralba, 2006). Gist can be extracted in the absence of fine detail, from degraded scenes, or when objects are difficult to process (Larson & Loschky, 2009; Meng & Potter, 2008; Potter, 1975; Potter, 1976; Oliva & Torralba, 2007; Rousselet, Joubert, & Fabre-Thorpe, 2005; Torralba, 2009). Oliva (2005) stated that gist occurs in two forms: conceptual and perceptual gist. Perceptual gist represents the depiction of the scene as defined by scene's global features, while conceptual gist is the semantic meaning of

the scene that is extracted during the cognitive processing that occurs after the viewing of the scene. Conceptual gist is influenced by perceptual gist.

Seminal research regarding gist found that when prompted with a verbal description or image of a target, participants were capable of detecting targeted scenes above chance during RSVP tasks when the scenes were displayed for less than a fixation (Potter, 1975; 1976). The extraction of gist has been shown to occur even when multiple scenes were presented on the screen simultaneously (Potter & Fox, 2009). Additionally, the type of prompt displayed indicating targets affected performance where pictures prompts resulted in higher detection performance than did text descriptors (Potter, 1976). Finally, when the prompts have more detailed or richer information, (i.e., butterfly vs. animal), perception performance typically increases (Intraub, 1981).

Researchers found that the longer scenes were displayed, the more rich the descriptions of the scenes or the features detected (Fei-Fei, et al., 2007; Intraub, 1981; Loftus, Nelson, & Kallman, 1983). Fei-Fei, et al. (2007) found a rich variety of information could be derived from a scene in 107 ms, which ranged from objects to scene classifications. Other research has shown that objects could be decoded as quickly as 100 ms (Lui, Agam, Madsen, and Kreiman, 2009).

Extraction of semantic information from scenes has been found to be robust even during scene degradation. Scenes have been detected and recognized when they were partially occluded (Meng & Potter, 2008), inverted (Diamond & Carey, 1986; Epstein, Higgins, Parker, Aguirre, & Cooperman, 2005; Evans & Treisman, 2005; Harding & Bloj, 2010; Kelley, Chun, & Chua, 2003; Meng & Potter, 2008; Shore & Klien, 2000), had color removed (Meng & Potter, 2008; Rousselet, et al., 2005), contained object inconsistencies (Biederman, Mezzanotte, &

Rabinowitz, 1982; Davenport, 2007; Davenport & Potter, 2004), and where the scenes were low resolution or poor quality (Loschky, Hansen, Sethi, & Pydimarri, 2010; Oliva & Schyns, 1997; Torralba, 2009). The recognition and detection of scenes in such scenarios has been attributed to the semantic information derived from the scene.

During such studies exploring gist, viewers were typically displayed a prompt, followed by a stimulus for durations that lasted up to a few hundred milliseconds, which was then followed by mask. Masks were used to cease perceptual processing of the stimulus (Potter, 1976). After the mask was been presented, the viewers were then asked whether the stimulus matched the prompt shown at the beginning of the task. Another approach has been to display a prompt, followed by displaying serially a rapid series of stimuli (RSVP) and asking if one of the presented stimuli matched the prompt shown at the beginning of the trial. The rapid series of stimuli effectively halts perceptual and conceptual processing of the previously presented stimuli (Intraub, 1984; Potter, 1976). In either task, participants typically achieved above chance performance for detecting the presence of targets provided by the beginning prompt. Typically, the stimuli were displayed from 10s of milliseconds to a few 100 milliseconds.

While longer presentation times may lead to richer descriptions, it may be of interest to limit the exposure to the stimulus to less than a fixation. This can be accomplished through the use of interstimulus intervals (ISIs), which remove the stimuli from the screen without masking and allow for continued processing. Loftus, Shimamura, and Johnson (1985), noted that performance on unmasked stimuli was equivalent to approximately an additional 100 ms of exposure time, due to the prolonged sensory presence of the stimuli in the visual icon (Sperling,

1960; Neisser, 1967). Potter & Fox (2009) found that the usage of ISIs resulted in a decrease in performance, but the performance was still comparable to not using an ISI.

Much of the scene perception literature has relied on classifying scenes into superordinate and basic-level categories. In order to work within the framework provided by the scene perception literature, it is essential to explore how websites are categorized. Essentially, web pages are complex documents, consisting of a variety of elements arranged spatially within a single document. Some previous classification attempts have relied on groups of elements or the type of elements found within a web page, sometimes in combination with previous personal experiences (Crowston & Williams, 2000; Dillon & Gushrowski, 2000) while other attempts have focused on automation and examining the hierarchy and the occurrences of types of text within a web page (Rehm, 2002; Santini, 2006). These methods have typically resulted in genres created by researchers or automation, but not by users.

Researchers have noted that the web does evolve over time, which has interesting implications for classifications. Santini (2007) noted that some types of websites may be emerging or may just be unknown. For instance, before blogs were a mainstay on the Internet, they were considered an emerging genre. Similarly, Crowston & Williams (2000) found a large portion of their genres as being previously unknown. Both Santini and Crowston & Williams stated that web pages may be classified into multiple genres. One such example is the distinction between social networking and blogs. Market research from NM Incite found that of the largest social networking websites, three were actually blogs (Nielsen, 2012). These included Blogger, WordPress, and Tumblr. One study had participants examine home pages of the individual states in the United States and had users place them into groups and then

examined them in terms of form, function, and content over time (Ryan, Field, & Olfman, 2002). They noted the shifting importance of these dimensions over time. However, none of these classification methods address whether websites can be classified into a similar taxonomy as scenes.

CHAPTER 2

RESEARCH STUDIES

The application of scene perception and scene gist to web pages has yet to be extensively explored. For instance, what information can be accessed for websites within a glance and how quickly can it be accessed? While scene perception has been studied extensively for natural scenes, little research has examined what semantic information can be extracted from brief glances of a web page. Prior studies regarding brief exposures on web pages primarily focused on visual appeal, trust, and perceived usability (Albert, et al., 2009; Lindgaard et al., 2006; Lindgaard et al., 2011; Tuch, et al., 2012). The following research explores whether web pages can be classified into basic-level categorizations (i.e., news, search, shopping, social networks, blogs, etc.) similar to scenes, and whether these categorizations are sufficient for detecting web pages when they are displayed for less than a fixation.

Study 1 – Classification Study

While previous research has shown that web pages can be classified with a wide variety of methods (Crowston & Williams, 2000; Dillon & Gushrowski, 2000; Rehm, 2002; Santini, 2006), research into classifying web pages into basic-level categorizations similar to scenes has not been completed before. On the Internet, websites can fall into several different genres. In this study, several genres were targeted because users have likely experienced these types of websites in the past and they also represent some of the oldest and largest genres on the Internet. The categories included news, search, shopping (i.e., ecommerce), social networking, blogs, mapping websites, and corporate websites.

News websites (i.e., CNN.com, FoxNews.com, etc.) predominantly aggregate and display news stories to visitors. News websites represented a large segment of online visitors and is one of the oldest types of websites on the Internet. In 1991, the first online news was published by the Weekend City Press Review (Weekend City Press Review, 2013). In 2011, United States news oriented websites were accessed on average by 110,224,700 monthly visitors (Edmonds, Guskin, Rosenstiel, & Mitchell, 2012).

Search websites (i.e., Google.com, Bing.com, etc.) offer functionality to search indexed information found on the Internet. Search websites can be considered as one of the oldest and most used types of websites on the Internet (McBryan, 1994). As of January 2013, search engines accounted for 3 of the top 5 and 11 of the top 25 websites on the Internet by visitor traffic (Alexa, 2013).

Ecommerce websites (i.e., Amazon.com, Zappos.com, etc.) provide functionality and allow users to purchase products or services through the website. Ecommerce is one of the largest and oldest segments of websites on the Internet. Ecommerce predated the modern, commercialized Internet. The first known online shopping transactions occurred in the 1980s (Feldman, 2011). However, the first known transaction on the commercial Internet was an order for a pizza through Pizza Hut in 1994 (Webley, 2010). In 2011 in the United States, $194.3 billion of goods were sold through ecommerce websites (eMarketer, 2012). In 2012, sales on Black Friday, traditionally the biggest shopping day of the year, surpassed $1 billion in sales in the United States (comScore, 2012). Ecommerce accounted for 3 of the top 25 most trafficked websites on the Internet (Alexa, 2013).

Social networking websites (i.e., Facebook.com, Twitter.com, etc.) allow functionality to meet, communicate, and share with others over the Internet based on interests or activities. Social networking is a relative newcomer. The first recognized social networking website was Six Degrees, which was launched in 1997 (Boyd & Ellison, 2007). Social networking websites accounted for 4 of the top 25 websites on the Internet (Alexa, 2013). The largest social networking websites have hundreds of millions of users. In 2012, Facebook surpassed 1 billion active users, Twitter announced that the service had more than 200 million active monthly users, and Google Plus had 135 million active users (Facebook, n.d.; Gundotra, 2012; Wang, 2012). Nielsen (2012) stated that in 2009, social networks/blogs became the top online destination for users, which accounted for 9.2% of time spent on the Internet.

Blogs (i.e., WordPress.com, Tumblr.com, etc.) are generally considered a form of a personal journal (Blog, n.d.). The first known blog has not been confirmed, but it was thought to have been created in 1997 (McCullagh & Broache, 2007). In 2011, the number of blogs on the Internet surpassed 181 million (Nielsen, 2012).

Map websites (i.e., maps.google.com, Mapquest.com, etc.) offer users the ability to view a map, locate information on a map, and find directions to destinations. MapQuest is likely the first online mapping website, which was founded in 1996 (Britannica, 2012).

Corporate websites (i.e., Honeywell.com, etc.) provide information about their respective companies. Corporate websites were of interest in this study because many of the websites that can fall into the categories in this study are representative of the companies themselves. This provides a website genre that potentially conflicts with the other types of website genres.

General knowledge websites (i.e., Wikipedia.org) were also included and were considered more general than the other categories used for classification. This category was intended to serve as a catchall for websites that look like they were used to present general information to visitors.

The purpose of this study was to classify web pages into conceptual categorizations based upon the utility and purpose of the websites. The results of this study provided categorized stimuli that were used in the two subsequent studies.

Method

Participants

For web page classification, 271 participants, all 18 years of age and older, were recruited from Wichita State University, the Wichita community, and Amazon's Mechanical Turk. Participants volunteered for the study through SONA Systems, which is the University experiment management system, by contacting the researcher directly, and through Mechanical Turk. Participants volunteering via SONA Systems were compensated by course credit. Participants volunteering via Mechanical Turk were compensated at an approximate rate of $6.00 per hour based on completing 50 classifications every 10 minutes. This rate was suggested by Amazon (Amazon, 2012). The actual compensation rate per hour varied depending how quickly they accomplished the classifications. Nine participants were removed due to poor classifications leaving data for 262 participants (171 females, 91 males; *M* = 26.05 years, *SD* = 10.63 years) for subsequent analysis.

Materials

Software. An online tool was developed for participants to use in order to categorize web pages. The tool displayed a consent form for the study, presented a short demographics survey, displayed instructions for completing the classification, preloaded stimuli prior to classification tasks, and collected categorization data. The online tool displayed the web pages at 512 px to the participants. Users accessed the tool via a web browser. Unless a participant accessed the study via SONA Systems, no personal identifying information was associated with demographic information. The participants' SONA Systems identification numbers were only used to grant SONA credit to the participants. The identification numbers were not used in conjunction with participant classification data to grant credit.

Categories. The website categories included news websites, search websites, shopping websites, social networks, blogs, maps, corporate websites, and general knowledge websites. Participants were also allowed to choose "None" when they felt the web page sample did not fit one of the provided classifications.

Stimuli. Websites (n = 132) were selected for classification across the categories. The websites selected did not represent an exhaustive list, but rather were websites that were considered characteristic of the categories in the study. The number of exemplar websites varied per category. Selected websites primarily targeted news, social networks, shopping, search, blogs, and maps. For both corporate and general knowledge websites, no specific websites were selected because it was thought that the websites selected for other categories could possibly be classified as these two categories. Websites were selected from Alexa's most

trafficked websites (Alexa, 2013), Google search results, and other categorical website lists/directories. See Appendix A, Tables 32-36 for a list of the websites used in the studies.

For news, 33 websites were selected. These ranged from local to national and international news websites. For shopping, 35 websites were selected. Shopping websites included ecommerce sites for manufacturers that sold a small product line to manufacturers that offered several hundred product lines. Additionally, ecommerce websites were included that sold 3rd party products from a few to many different manufacturers.

For search, 34 websites were selected. Search websites were primarily limited to websites that either presented a primary search page to enter search terms or presented predominantly textually-oriented search results. Services that provide image or video search results were excluded from the sampling.

For social networking, 28 websites were selected. Fewer exemplar social networking websites were available than the other categories. The current landscape of social networking on the Internet oriented toward English speakers has typically been dominated by relatively few large services such as Facebook, Twitter, and Google+. There were some large social networking websites targeted toward non-English speakers, but these social networks were excluded from this sample because they did not provide English language versions of their services, or the content was in another language other than English. Several social networking websites required users to pay to access or communicate with other users. These social networking websites that limited functionality due to a pay-to-interact model were excluded from the sampling. Generally, websites designed exclusively for dating were excluded from this sampling, even though they may be considered social networks. Some social networking

websites promoted dating as a secondary function through their service, but were not excluded. The social networking sites included in the sample were primarily oriented toward casual and business social networking, but also included music, games, sharing information, and photography.

An additional 19 map-oriented websites were included to match the mapping category. Map-oriented websites included services that provided maps, transportation directions, weather, and historical maps.

For blogs, stimuli for both news and social networks could possibly be categorized as blogs. News websites were sometimes published using blogging platforms or published in a blog format. Moreover, blogs have been noted to be considered as social networks according to market research, so it may be possible that these would be considered as blogs.

Screenshots were captured for the websites at 1,280 x 1,024 px, which effectively left 1,250 x 972 px of usable web page real estate after the browser user interface was removed from the screenshot.

All screenshots of the web pages were debranded and had extraneous, identifying décor blurred with a Gaussian blur in Adobe Photoshop and TechSmith SnagIt Editor. The blur was large enough to completely cover the element. This essentially removed the element, but retained some visual information that an element was in the location, thus preserving some spatial integrity provided by the element. In some cases, the amount of blurring essentially blanked out the element. Extraneous elements included labeling that served no other purpose than to inform the user of the purpose of the website. These included taglines, slogans, and so

forth. To protect the privacy of users on social networking websites, names and faces were blurred. Across all websites, obscenities were also blurred.

The stimuli were divided randomly into multiple blocks. The number of stimuli in a single block was approximately 50 screenshots. Smaller blocks were made available to participants after the initial classification routine began, which consisted primarily of screenshots that contained variations of web pages with different advertisement configurations, screenshots of websites that had changes in layout from its initial classification, or screenshots of new websites added to the classification study.

Procedure

Participants accessed the study via a study link posted on SONA Systems, Facebook, Mechanical Turk, or in the Wichita area. Upon entering the classification tool, they were asked to read and confirm agreement or disagreement with a consent form. After providing consent, they were asked to complete a short demographics survey which included age and gender. After completing the short survey, the tool cached all stimuli into their browser to reduce the amount of loading between stimuli. When this was completed, the participant continued to the instruction page. After reading the instruction page and selecting next to continue, they were shown the categorization screen. The categorization screen displayed a screenshot of a web page at 512 x 386 px, 9 checkboxes with associated labels, and a button to continue. Participants classified a stimulus into one or more of the categories, which included news websites, shopping websites, search websites, social networking websites, blogs websites, maps websites, corporate websites, general knowledge websites, or none of the above. Upon finishing a classification, participants clicked a button to proceed to the next classification.

Participants were not limited in the amount of time they could spend examining each stimulus. See Figure 1 for a screenshot of the classification application.

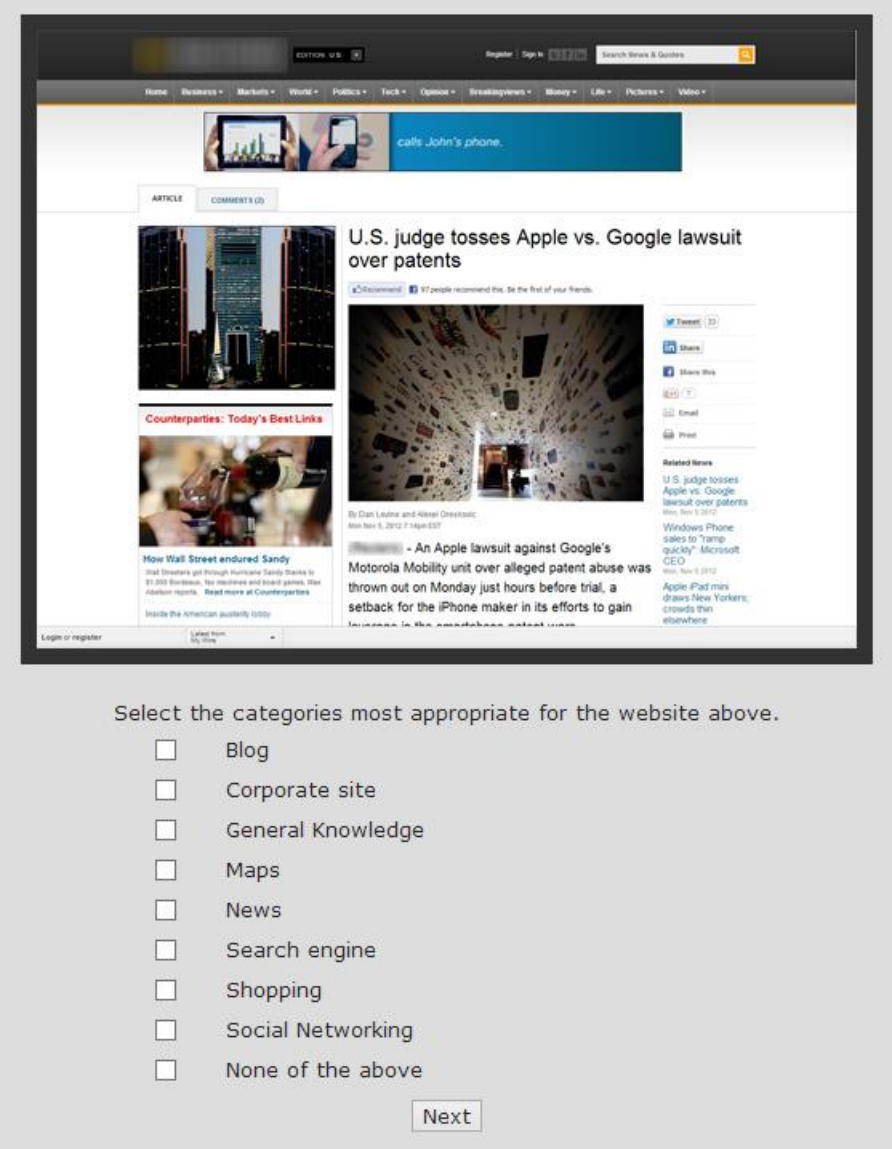

Figure 1. Screenshot of classification application.

Participants were allowed to classify multiple blocks of stimuli, but were prevented from classifying the same stimulus multiple times. Once a block was started, participants were not allowed to re-access the block at a later time in order to prevent duplicate classifications.

Results

On average, each participant classified approximately 46.23 stimuli (*SD* = 36.61). Each stimulus was classified into an average of 1.19 categories (*SD* = .33). The large standard deviation can be attributed to participants quitting early, and additional, smaller blocks, which were described in the procedure. Participants spent an average of 8.55 seconds (*SD* = 6.93 seconds) categorizing each screenshot. This equated to participants spending about 6 minutes and 5 seconds on average (*M*= 364.83 seconds; *SD* = 324.33 seconds) categorizing screenshots. Participants were screened for outliers. In total, nine participants were removed due to poor classifications. See Appendix B for more information.

Selection criteria was set where 80% of participants agreed that a stimulus was a particular classification, but no more than 20% of participants agreeing it was different type of classification.

After examining the classifications, many of social networking web pages were classified as both blogs and social networks. Nielsen (2012) noted the distinction between social networks and blogs may be blurred. They found that three of the largest social networking websites were actually blogs. In this classification study, the distinction between blogs and social networks was not strong. Due to this, blogs and social networks were collapsed into a blogs/social networks category. The same criteria were set for this new classification group (80%/20%).

The agreement level for each of the four website categories was calculated. For each stimulus in a category, its level of agreement was equal to the percentage of participants that agreed its exemplar belonged to the respective category. The exemplars were weighted

according to the number of times they were represented in each category and each category had 276 screenshots. As such, analyses for both were conducted. See Figure 2 and Table 1 for website category agreement.

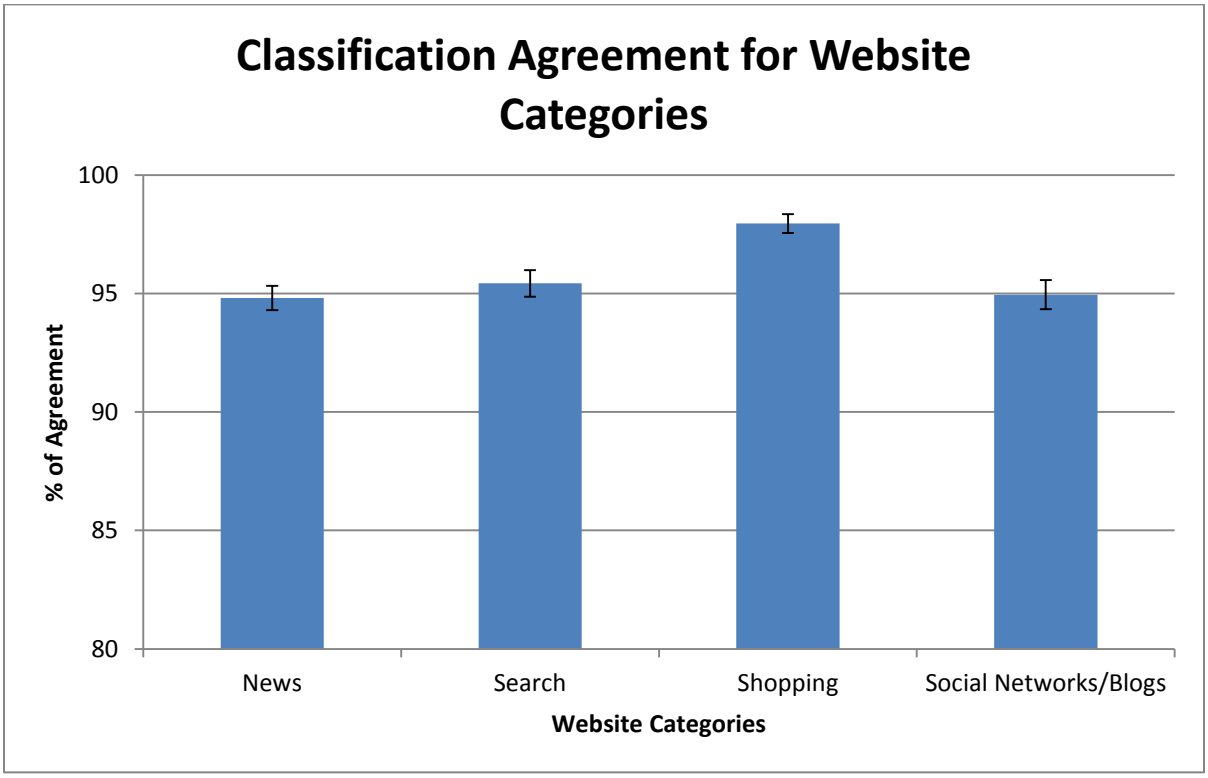

Figure 2. Agreement across the website categories.
Note: Error bars are 95% confidence intervals

TABLE 1

PARTICIPANT AGREEMENT FOR WEBSITE CLASSIFICATION

Website Category	N	Overall Rank	Agreement (*SD*)
Search Websites	276	3	95.43% (4.75%)
Shopping Websites	276	1	97.95% (3.33%)
Social Networks/Blogs	276	3	94.95% (5.16%)
News Websites	276	3	94.81% (4.33%)

The normality of the data for the four categories was examined. Skewness and kurtosis for measuring normality are sensitive to large sample sizes (Field, 2009; Tabachnick & Fidell, 2007). Guidelines have been proposed to help to evaluate skewness and kurtosis for medium to large samples (Breakwell, Hammond, Fife-Schaw, & Smith, 2006). If the assumption that skewness and kurtosis fall within the range of ±1, these may be considered close to normal. The Shopping category fell outside both of these criteria (skewness = -2.14, kurtosis = 5.64). Evaluation of histograms and Q-Q plots confirmed negative skewness was across each category. Shapiro-Wilk tests indicated each of the four categories did not meet normality assumptions, $p < .01$. The data was not transformable to normal distributions through transformations. Due to the lack of improvement regarding overall normality, data was further analyzed with nonparametric statistics.

To determine the differences in the level of agreement between the four categories, a Kruskal-Wallis test was conducted. Results indicated significant differences between the four categories, $H(3) = 89.21$, $p < .01$. Post-hoc Mann-Whitney *U* tests with Holm-Bonferroni corrections were conducted. Field(2009) suggested calculating *r* as a measure of effect size for Whitney U, Wilcoxon, and dependent samples *t*-tests, which follows the following interpretation: $r < .3$ = small; $.30 <= r < .5$ = medium; $r >= .50$ = large. For r^2, this translates into $r^2 < .09$ = small; $.09 <= r^2 < .25$ = medium; $r^2 >= .25$ = large. Results indicated that shopping had higher agreement than social networks/blogs ($U = 25500.5$, $r^2 = .10$, $p < .01$), news ($U = 21735.5$, $r^2 = .15$, $p < .01$), and search ($U = 26410.5$, $r^2 = .08$, $p < .01$). Overall, differences between shopping and the other categories had small to medium effect sizes. However, no differences were noted between search and news ($U = 34667.5$, $r^2 = .01$, $p = .06$), search and social

networks/blogs (U = 36448, r^2 < .01, p = .37), and social networks/blogs and news categories (U = 36629, r^2 < .01, p = .43).

Discussion

The results of this analysis suggest that shopping websites have the greatest agreement amongst the four categories, with no differences between search, news, and social networks/blogs.

Study 2 – Perception of Scenes and Websites

Study 1 demonstrated that web pages can be classified into basic-level website categories. Several hundred web pages were classified into nine different categories based on their perceived functionality, which included blogs, social networks, search, news, shopping, mapping, general knowledge, corporate websites, or none of these categories. In all, the majority of the web pages were classified with at least 80% agreement that the web page was one type of category and no more than 20% agreement that it was another type. Because these web pages had such strong categorizations, it may be plausible that web pages can be detected when displayed for less than a fixation based on this classification schema. One aspect of Dillon & Gushrowski (2000) was that home pages with similar agreement typically had similar web elements. Similar results were found in the classification study where web pages with similar classifications shared similar features.

A pilot study was conducted to determine the well how web pages may be discriminated when compared with scenes. The results of this initial pilot study indicated that participants could discriminate between both web pages and scenes, but the discriminability for web pages was lower.

Previous literature has indicated that judgments regarding attributes of web pages can be based on a mere glance. The classification study suggests that users have conceptualizations of websites, and the results of the pilot suggest users likely have schemas for web pages that can be accessed quickly. This raises an important question for how web pages are represented.

The purpose of this study was to determine whether web pages could be detected during RSVP tasks with above chance accuracy when displayed for less than a fixation. In order to provide baseline comparisons, participants were asked to detect natural scene targets in separate conditions. The natural scene conditions provided a "gold standard" condition for gist and scene-related performance. In contrast, a "lead standard" condition was also created to provide a comparison of degraded performance based on the interruption of extracting a scene's semantic meaning. In this study, the scenes for the "lead standard" condition were degraded through inverting the natural scenes.

The use of inverted scenes has several advantages. Features such as spatial structure, color, and luminance remain consistent regardless of orientation, but the change in orientation should interrupt detection based semantic knowledge of the scenes. The results of using inversion (180 degree rotation) as a method of degrading scenes have been mixed. Inversion has been supported to reduce performance in the detection of a scene target during RSVP tasks (Evans & Treisman, 2005), for faces more so than landscapes (Diamond & Carey, 1986), but not for detecting animals and humans (Rousselet, et al., 2003). When detecting changes, inversion was found to decrease performance (Kelley, et al., 2003; Shore & Klien, 2000). Inversion was shown to reduce performance when used in combination with occlusion (Meng & Potter, 2008) and changes in luminance (Harding & Bloj, 2010), but not significantly when in conjunction with

grayscaling (Nandakumar & Malik, 2009) and jumbling scenes (Zimmermann, Schnier, & Lappe, 2010).

Previous research indicated that action video game play may lead to increased spatial cognition (Feng, Spence, & Pratt, 2007), increased visual attention (Green & Bavelier, 2003; Green & Bavelier, 2007; Dye, Green, & Bavelier, 2009). However, other research has had difficulty replicating such results (Murphy, & Spencer, 2009). Given this, gaming behavior was of interest in the study as it may result in better performance.

In this study, stimulus onset asynchrony (SOAs) durations were measured. A SOA consists of the presentation time and the ISI. SOAs were measured at several desired accuracy thresholds for detecting upright scene, inverted scene, and web page targets during rapid serial visual presentation (RSVP) tasks.

Several hypotheses were developed based on previous literature, the classification study, and initial pilot testing.

H_1: Participants will be able to discriminate between categories for both scenes and web pages based on stimulus exposure lasting less than a fixation, but participants will have worse performance for web pages than for scenes. Overall, as the accuracy threshold increases, the SOA should also increase for both scenes and web pages.

H_2: The necessary performance to reach desired accuracy thresholds will be lower for categories with higher agreement. Shopping web pages should be more readily discriminated than the other website categories.

H_3: Performance will be better for upright scenes than inverted scenes, but inverted scenes will be similar or better than web page related SOAs. From the pilot testing, upright

scenes had better discriminability than websites. Moreover, literature has provided that the inverting of scenes may result in longer SOAs for the inverted condition. As such, SOAs for websites and inverted scenes may be similar or differences may be found.

Method

Participants

Twenty-five participants, 18 years of age and older, were recruited from Wichita State University and the Wichita Metropolitan Area. Participants volunteered for the study through SONA Systems, the University experiment management system, or by contacting the researcher directly. Participants volunteering via SONA Systems were compensated with course credit. Out of the 25 participants, six participants were subsequently removed. Two participants did not complete all sessions of the study. Three participants were screened due to suboptimal near visual acuity. One participant was excluded from subsequent analysis due to poor overall performance. Overall, 19 participants (7 males, 12 females; *M* = 21.16 years, *SD* = 3.67 years) completed the study and were kept for subsequent analysis. See Appendix C for more details. For the remaining participants, additional demographic information can be found in Appendix D.

Materials

Equipment. The experiment was conducted using a 2GHz Apple Mac Pro computer driving a 22-inch CRT monitor. The CRT monitor had a refresh rate of 85 Hz and a resolution of the monitor was 1400 x 1050 px. A chinrest was used to maintain a consistent distance between the monitor and the participant during the experiment.

Matlab and RSVP software written by the researcher was used in the experiment. The software incorporated functionality of PsychToolbox (Brainard, 1997; Pelli, 1997; Kleiner, Brainard, & Pelli., 2007) and QUEST (Watson & Pelli, 1983; Pelli, 1987).

Questionnaire. A questionnaire was used in the experiment to gather information regarding demographics, computer usage, web usage, and video gaming experience. Demographic questions included age, gender, student status and major (if applicable), and occupation. The contents of the questionnaire can be found in Appendix E.

From Owens, et al. (2011), three questions regarding computer and Internet usage were included in the questionnaire. Participants were asked how often they used a computer and the Internet on a weekly basis. In addition, they were asked to select why they use the Internet. Additional questions were developed inquiring about the usage of news websites, search engines, shopping websites, and social networks/blogs. Finally, they were asked which websites they use frequently for the respective category.

Questions regarding video game experience were included in the study as well. Two questions regarding a participant's current and maximum frequency of video game play per week were included from the Gaming Experience Measure (GEM) (Taylor, Singer, & Jerome, 2008). Finally, two questions regarding frequency of play and skill regarding 14 video game genres was included (Phan, Jardina, Hoyle, & Chaparro, 2012). While a similar set of genres was used in the GEM, the genre-related questions from Phan, et al. included current game genres like social network games.

Stimuli. Four website categories were selected for the study, which including search, news, shopping, and social networks/blogs. The website stimuli were selected and validated in

the classification study. For news, shopping, and search, 276 screenshots were used. For social networks/blogs, 172 screenshots were used. See figure 3 for examples of the website stimuli.

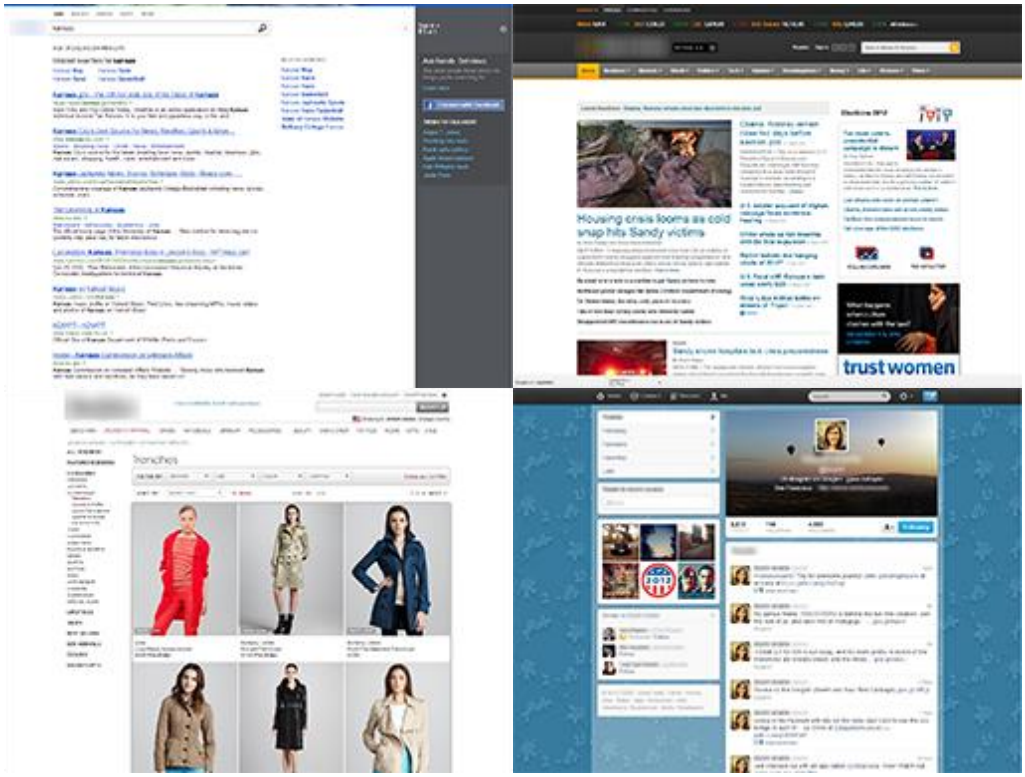

Figure 3. Examples of website stimuli used in the study.
Note: Clockwise from upper left: search, news, social networks/blogs, shopping.

Four basic categories of scenes were selected for the study, which included beaches, mountains, deserts, and forests. The four scene categories were related on a superordinate level. Using the schema provided by the SUN database (Xiao, Hays, Ehinger, Oliva, & Torralba, 2010) the scene categories were representative of outdoor natural scenes. For this experiment, beaches and mountains were used as target categories and deserts and forests were used as distractor categories. Additional inverted versions of the scenes were created.

Images were selected from the SUN database (Xiao, et al., 2010) and also using image search engines, such as Google Image Search. Scene stimuli used in the experiment was

validated through pilot testing. For each of the target categories, 268 scenes were selected. For each distractor category, 284 scenes were selected. See figure 4 for examples of natural scenes.

Figure 4. Examples of natural scene stimuli used in the study.
Note: Clockwise from upper left: desert, beach, forest, mountains.

Both web and scene stimuli were displayed at 512 x 386 px, which was a scaled version of web pages displayed in a web browser at 1,280 x 1,024 px with the vertical scrollbar and browser UI removed. At 60cm, the stimulus subtended 13.69 degrees horizontally and 10.34 degrees vertically.

Procedure

Participants were seated at a desk in the laboratory and they were asked to read and complete a consent form. After they provided consent by signature, the facilitator conducted a

near visual acuity test using a Lighthouse Near Acuity Chart and a color vision test using Dvorine Pseudo-Isochromatic Plates (Dvorine, 1963).

Once both of the vision tests were completed, the facilitator instructed the participant to read the procedure and descriptions of the stimuli for the experiment aloud. The procedure and descriptions can be found in Appendix F. The facilitator reviewed the stimuli descriptions with the participant, asking them to describe each of the different types of stimuli as they interpreted it. This process helped eliminate confusion across the different stimuli categories. After the discussion regarding the stimuli, the procedure was reviewed with the participants, and then they were asked if they had any questions regarding the study. Once questions were answered, the participant was seated at the experiment computer in front of a chinrest 60cm from the monitor and the facilitator entered the parameters for the experiment into the experiment software.

Instructions for keyboard commands were displayed on the screen for how participants should complete the trials. Participants were instructed to respond to trials where they thought the target was present by pressing the “p” button. Similarly, if they thought the target was absent, they were instructed to respond with the “q” button. The spacebar was used to begin the trials and to continue after the built-in experimental breaks.

The software prompted participants to press the spacebar to begin when ready. Each trial consisted of the display of a word description (target stimuli category), a fixation point, several stimuli presented in rapid succession, and a response prompt. Participants were instructed to respond to whether the target category was present within the RSVP stream.

The verbal description of the target was displayed on the screen for 2,000 ms, followed by a fixation dot. The text displayed on the screen stated "Target: trial category" where trial category was the target for that respective trial. A fixation dot remained on the screen for 300 ms, and then the screen was blanked for 200 ms, after which the RSVP stream of stimuli was displayed. The response prompt appeared immediately after the RSVP stream for the participants to respond and remained on the screen until the participant responded or 5 seconds elapsed. If the participant did not respond within that time frame, the trial was marked with no response and incorrect. No feedback was given to the participant regarding whether their response was correct or incorrect. See Figure 5 for the schematic of a RSVP trial.

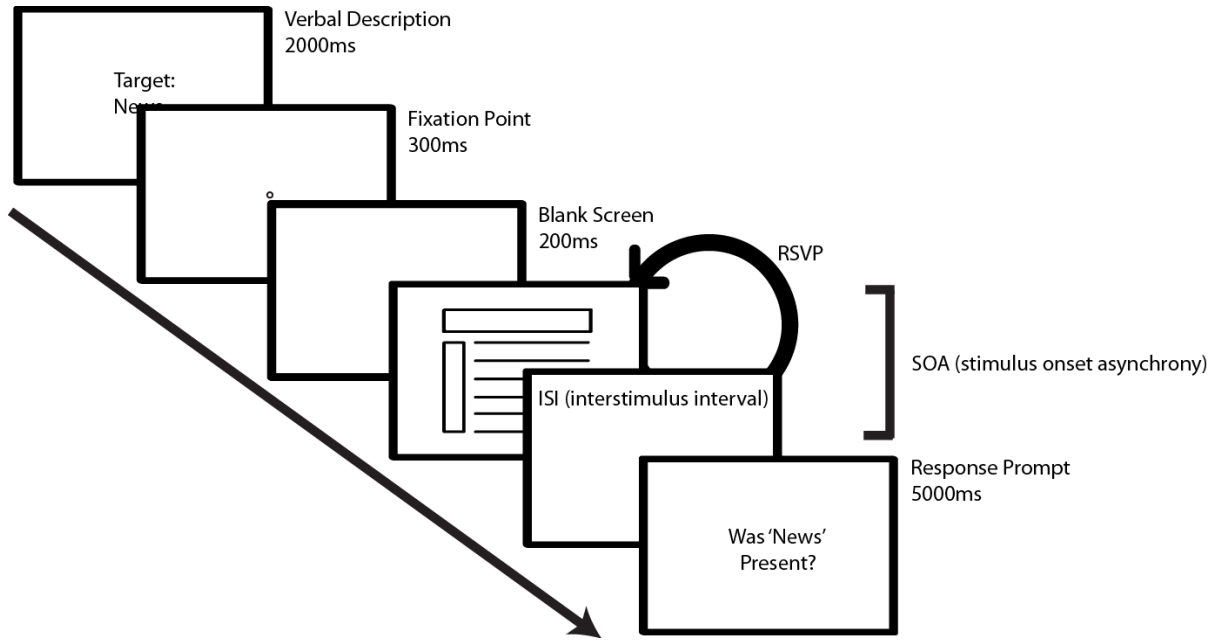

Figure 5. RSVP trial schematic for study 2.

In the RSVP stream, each frame had a stimulus onset asynchrony (SOA). This is the duration from the onset of a stimulus to its masking, which consisted of presenting the stimulus for a particular duration and an interstimulus interval (ISI). This limited each stimulus in the

stream to a single fixation. Given this, the maximum presentation time during the RSVP trial was 140 ms, which is less than a typical fixation during a scene perception task. The display of the stimuli was followed by an ISI if the SOA exceeded the presentation time. During the ISI, the stimulus was removed from the screen and replaced with the background color, which allowed for continued processing of the icon and additional processing time until the next stimulus was displayed. For each RSVP stream, the SOA remained consistent within a respective RSVP stream, but varied between streams.

The SOA for each RSVP stream was determined using the QUEST algorithm, a frequently used, statistical-based adaptive staircase method (Watson & Pelli, 1983). QUEST is a Bayesian method of determining the threshold required to achieve a predetermined level of performance. Since participants were responding to whether a target was present or not, 50% represented chance performance. QUEST sought estimates of the SOAs for above chance performance for both website and scene categories at three accuracy thresholds: 60%, 75%, and 90% correct.

QUEST required estimated parameters for the mean and standard deviation for a desired threshold (Watson & Pelli, 1983). Through pilot testing, it was determined that SOAs should be limited to a defined range. The minimum SOA was limited to a single screen refresh at 85hz. The maximum SOA was equal to one second. When using a range in QUEST, the QUEST procedure used the median of the range as the initial mean (505.9 ms). For the standard deviation, one second was used.

Twelve QUEST staircases were created for websites and scenes. One staircase was created for each type of website. News, search, shopping, and social networking/blogs were

tested at the three defined accuracy thresholds (e.g., 60%, 75%, and 90%). Similarly, staircases for mountains and beaches were created for each type of scene (e.g., upright and inverted). The mountain and beach target scenes categories were tested at the same defined accuracy thresholds.

For each category/accuracy combination, participants completed 2 practice and 40 experimental trials. Overall, participants completed 24 practice trials at the beginning of the experiment followed by 480 experimental trials for the website session. The same was repeated for the scene session. The trials were randomized.

For both websites and scenes, participants performed 252 target-present and 252 target-absent trials. Target-absent or target-present trials were presented in a random order. For each trial, 15 stimuli, five from each nontarget category, were selected randomly. Within each trial, the selection was without replacement. However, between trials, each nontarget stimulus may have appeared in other RSVP streams. During a target-present trial, the RSVP stream contained one stimulus from a target category that was selected without replacement. It was placed randomly, but neither in the first nor last position in the RSVP stream. As such, the participant saw targets only once, while nontarget stimuli were seen multiple times.

If participants exhibited the inability to perform the task for a staircase, the staircase was ceased and the remaining trials in the staircase were skipped. The criterion for skipping trials was the QUEST mean exceeding the specified maximum range minus two screen refreshes for 10 consecutive trials. This equated to QUEST means greater than 976.5 ms. The likelihood of the QUEST mean improving below this threshold was minimal, given the QUEST standard deviation due to small standard deviations. If the QUEST mean dropped below this threshold

prior to the staircase stopping, the counter was reset. This procedure prevented participants from completing unnecessary trials and to help maintain that the experiment lasted no longer than 120 minutes.

Website and scene stimuli were tested in separate experimental sessions. The testing order for website and scene sessions was counter-balanced to eliminate order effects. Total duration of the experiment lasted approximately 120 minutes per experimental session. Participants were instructed to take 1-2 minute breaks after every 15 minutes of RSVP trials. All sessions were monitored to ensure that participants were on task and using the chinrest.

After completion of each session, participants were queried about the strategies they deployed to detect target categories. These responses were recorded. After the completion of both sessions, participants completed the post-experiment questionnaire and then were debriefed about the experiment.

Results

The data was checked for normality assumptions. Details regarding decision criteria for transformations and statistical procedures can be found in Appendix G. Additional results for collapsing across groups, web usage, and video game play can be found in Appendix H.

QUEST mean estimates of SOAs

The means of the QUEST estimates for SOAs for the 60%, 75%, and 90% accuracy thresholds for each type of web and scene stimuli were calculated. In this study, the SOA for each category at each accuracy threshold was the presentation time (plus interstimulus interval, if applicable) required to detect a stimulus matching a target category during RSVP

tasks. The mean SOAs for stimuli types (upright scenes, inverted scenes, and website) can be found in Table 2 and figure 6.

TABLE 2

QUEST MEAN SOAS FOR EACH CATEGORY ACROSS ACCURACY THRESHOLDS.

	QUEST Mean Estimates of SOAs in milliseconds (*SD*)			
Category	60%	75%	90%	Overall
Websites	328 (182)	594 (232)	824 (143)	582 (168)
Upright Scenes	109 (28)	248 (190)	539 (202)	299 (123)
Inverted Scenes	198 (128)	383 (244)	593 (212)	391 (170)
News	397 (291)	641 (338)	840 (228)	626 (204)
Search	272 (235)	516 (343)	805 (200)	531 (220)
Shopping	257 (239)	449 (296)	698 (312)	468 (231)
SN/Blogs	387 (234)	773 (240)	953 (74)	704 (150)
Upright Beaches	108 (47)	207 (173)	490 (330)	269 (156)
Upright Mountains	109 (44)	289 (256)	588 (303)	329 (145)
Inverted Beaches	214 (165)	380 (263)	583 (245)	392 (157)
Inverted Mountains	183 (173)	386 (279)	602 (290)	391 (219)

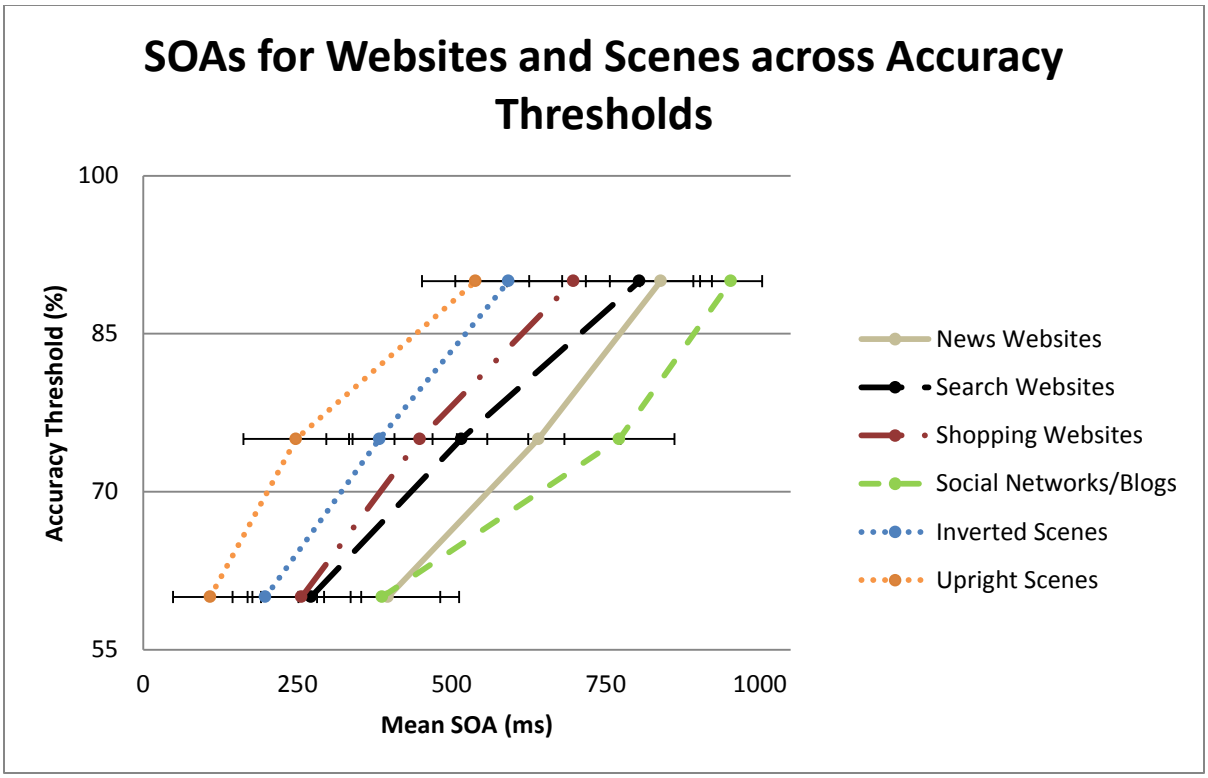

Figure 6. Mean SOAs for websites and scene categories accuracy thresholds.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

Poor performance

The number of skipped trials was a good indicator for poor performance during the study. Primarily, participants skipped trials for news websites, social networks/blogs, search websites, and shopping websites at the 90% accuracy threshold. These data points equated to poor performance during the experiment and were not treated as outliers. Removal of these data points would artificially lower the SOA for those respective categories and thresholds. Given this, the 90% threshold means for those website categories were treated as a lower limit for a respective mean SOA in this study. See Table 3 for the number of participants that skipped trials per condition.

Only 2 participants for the scene conditions skipped trials. One participant skipped trials in both the 75% upright mountain conditions and the 90% inverted mountain condition, while one other skipped trials in the 90% inverted mountain condition.

TABLE 3

MEAN NUMBER OF SKIPPED TRIALS AT 60%, 75%, AND 90% ACCURACY THRESHOLDS.

	60%		75%		90%	
Category	N	M (SD)	N	M (SD)	N	M (*SD*)
Websites	0	-	0	-	12	9.04 (*4.94*)
Upright Scenes	0	-	1	9 (-)	0	-
Inverted Scenes	0	-	0	-	2	5.5 (*.71*)
News	0	-	0	-	8	5.46 (*9.88)*
Search	0	-	0	-	3	3.24 (*7.00*)
Shopping	0	-	0	-	5	1.16 (*4.33*)
SN/Blogs	0	-	0	-	7	4.93 (*11.57*)
Beaches (U)	0	-	0	-	0	-
Mountains (U)	0	-	1	9 (-)	0	-
Beaches (I)	0	-	0	-	0	-
Mountains (I)	0	-	0	-	2	5.5 (*.71*)

Note: (U) indicates the upright scene condition. (I) indicates the inverted scene condition. Dashes indicate no data available.

Overall performance across stimuli types

To establish a high-level overview of how participants performed, the SOAs for stimuli type were collapsed across each respective category for each accuracy threshold (60%, 75%, and 90%). This provided a mean SOA for each of the three stimuli types (websites, upright scenes, inverted scenes). See Table 4 for QUEST mean estimates of SOAs for each stimuli type.

TABLE 4

MEAN SOAS FOR EACH COLLAPSED STIMULI TYPE ACROSS ACCURACY THRESHOLDS

	QUEST Mean Estimates of SOAs in milliseconds (*SD*)			
Stimuli Category	60%	75%	90%	Overall
Websites	328 (182)	594 (232)	824 (143)	582 (168)
Upright Scenes	109 (28)	248 (190)	539 (202)	299 (123)
Inverted Scenes	198 (128)	383 (243)	593 (212)	391 (170)

To gain a better understanding of how performance varied across the three accuracy thresholds, the SOAs for each threshold were analyzed. A 3x3 repeated-measures ANOVA was conducted on log transformed SOAs across the stimuli and the accuracy thresholds. The results indicated a main effect for stimuli, $F(2,36) = 38.32$, $p < .01$, partial $\eta^2 = .68$, a main effect for accuracy threshold, $F(2,36) = 197.05$, $p < .01$, partial $\eta^2 = .92$, and an interaction between stimuli type and accuracy threshold, $F(4,72) = 3.92$, $p < .01$, partial $\eta^2 = .18$. The interpretation of partial η^2 states that partial $\eta^2 = .01$ as being small, partial $\eta^2 = .06$ as being medium, and partial $\eta^2 = .14$ as being a large effect size (Stevens, 2009). Overall, the effect sizes for the main effects and for the interaction were all large. Planned pairwise comparisons were conducted to determine which conditions differed. For each accuracy threshold, upright scene and inverted scene SOAs were shorter than the SOA for websites. However, the SOAs for upright scenes and inverted scenes only differed at the 60% and 75% accuracy thresholds, but not for the 90% threshold. For the significant comparisons, the effect sizes were large. Across the three types of stimuli, increases in accuracy threshold resulted in significant increases in SOAs in order to discriminate targets, $p < .01$.

Figure 7 shows the mean SOAs of the three stimuli types. Table 5 reflects the back-transformed means of the log SOAs for each accuracy threshold used in the repeated measures ANOVA. See Table 6 for planned comparison statistical results and significant differences.

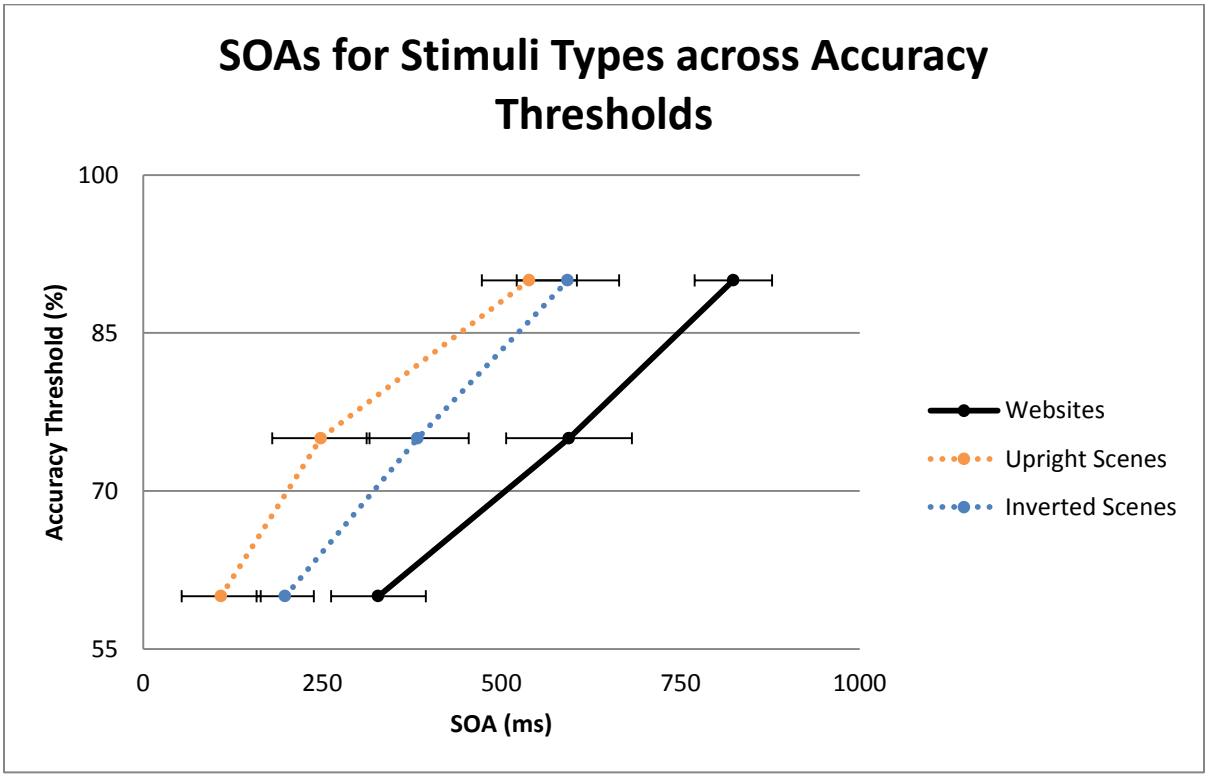

Figure 7. Mean SOAs for scene categories across accuracy thresholds.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

TABLE 5

MEAN SOAS FOR EACH COLLAPSED STIMULI TYPE ACCURACY THRESHOLDS

	QUEST Mean Estimates of log SOAs in milliseconds [*CI*]			
Stimuli Category	60%	75%	90%	Overall
Websites	291 [241, 352]	543 [447, 656]	812 [716, 922]	504 [431, 591]
Upright Scenes	105 [90, 123]	198 [157, 248]	498 [430, 574]	218 [195, 244]
Inverted Scenes	165 [138, 199]	321 [266, 387]	549 [468, 641]	308 [271, 348]

Note: Means are backtransformed. CIs are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008).

TABLE 6

STATISTICAL TESTS COMPARING WEBSITES, UPRIGHT SCENES, AND INVERTED SCENES

Comparison	*t*	*p*-value	r^2	MD (ms)
Websites - Upright Scenes (60%)	8.95	< .01*	.82	186
Websites - Upright Scenes (75%)	6.13	< .01*	.68	345
Websites - Upright Scenes (90%)	4.57	< .01*	.54	314
Websites - Inverted Scenes (60%)	4.24	< .01*	.50	126
Websites - Inverted Scenes (90%)	4.08	< .01*	.48	263
Upright Scenes - Inverted Scenes (75%)	-3.37	< .01*	.39	-124
Websites - Inverted Scenes (75%)	3.24	.01*	.37	222
Upright Scenes - Inverted Scenes (60%)	-3.04	.01*	.34	-60
Upright Scenes - Inverted Scenes (90%)	-.93	.36	.05	-51

Note: * indicates a significant comparison. MD is mean difference. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

At a high-level overview, participants required longer SOAs to detect website targets than targets from natural scene categories. Additionally, SOAs were shorter for upright scenes than for inverted scenes at 60% and 75% accuracy thresholds, but not at the 90% thresholds. Moreover, detection of inverted scenes required shorter SOAs than for between websites, regardless of accuracy threshold.

Performance across accuracy thresholds

SOAs for upright scenes, inverted scenes, and websites significantly differed from each other in the high-level overview. To better understand how performance changed across accuracy thresholds, the individual categories were compared. First, the four website categories were compared against each other.

To examine the main effect for the type of website, a repeated measures ANOVA was conducted on the overall square root transformed SOAs of each website category. The results

of the repeated measures ANOVA indicated significant differences between different types of websites, $F(3,54) = 12.20$, $p < .01$, partial $\eta^2 = .40$. Overall, the effect size for the differences between the different types of websites was large. Planned comparisons showed that SOAs associated with shopping and search websites were significantly shorter than SOAs for social networks/blogs. Similarly, the mean SOA for shopping websites was significantly shorter than the SOA for news websites. No significant differences were noted between news websites and social networks/blogs. See figure 8 for comparison of mean SOAs across the four categories. See Table 7 for back-transformed SOAs. See Table 8 for planned comparison statistical test results, backtransformed mean differences of the square root transformed SOAs, and significant differences.

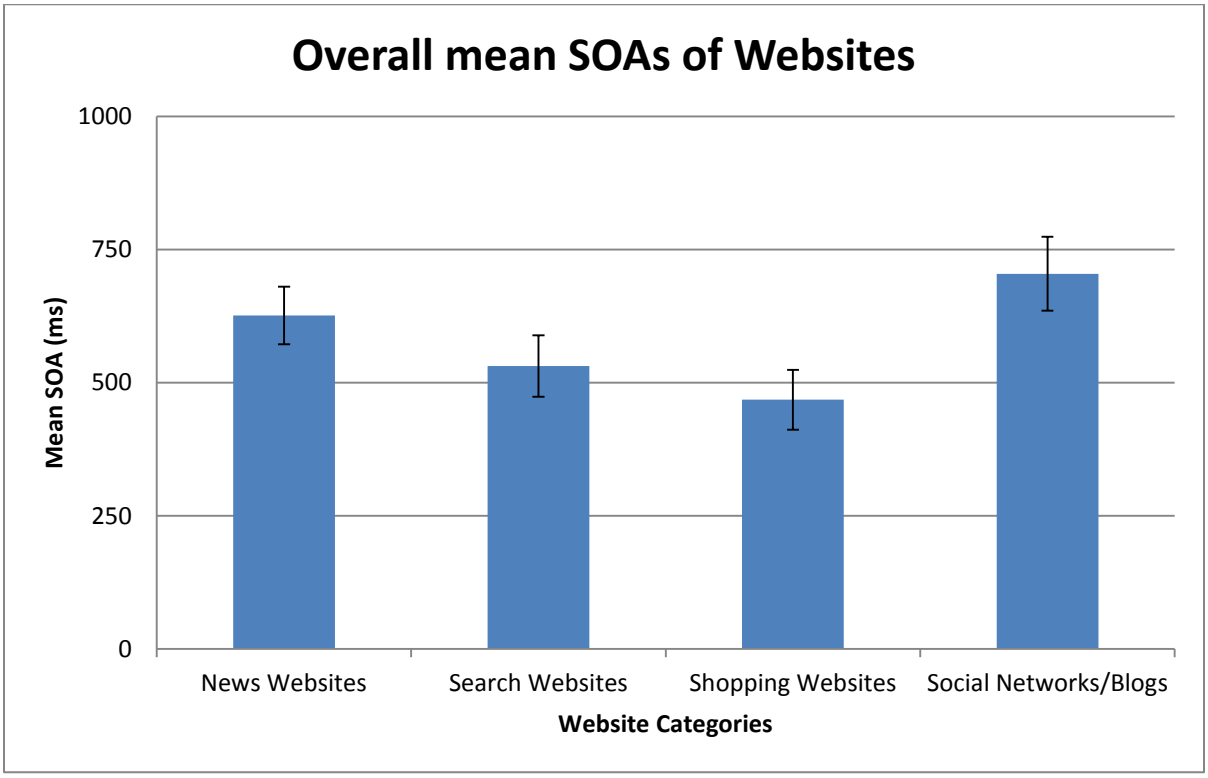

Figure 8. Mean SOAs across the four categories of websites.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

TABLE 7

MEAN SOAS FOR EACH WEBSITE CATEGORY COLLPASED ACROSS THRESHOLDS

Category	M [*CI*]
News Websites	610 [557–666]
Search Websites	509 [452–570]
Shopping Websites	439 [382–500]
Social Networks/Blogs	696 [620–777]

Note: Means are backtransformed. CIs are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008).

TABLE 8

STATISTICAL TESTS COMPARING OVERALL SOAS OF THE WEBSITE CATEGORIES

Comparison	*t*	*p*-value	r^2	MD (ms)
Shopping Websites - Social Networks/Blogs	-4.83	< .01*	.56	-257
News Websites - Shopping Websites	4.24	< .01*	.50	171
Search Websites - Social Networks/Blogs	-3.72	< .01*	.43	-187
News Websites - Search Websites	2.45	.03	.25	101
News Websites - Social Networks/Blogs	-1.87	.08	.16	-86
Search Websites - Shopping Websites	1.68	.11	.14	70

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

Next, the website categories were analyzed by accuracy threshold to determine how the SOAs differed. Friedman tests were conducted to determine the differences between websites. When examined by accuracy threshold, website SOAs were similar to each other at the 60% accuracy level, $X^2(3) = 7.67$, $p = .05$, $W = .14$. However, at 75% accuracy, SOAs significantly differed from each other, $X^2(3) = 9.38$, $p = .03$, $W = .17$. This trend continued for the 90% accuracy threshold where SOAs were significantly different from each other, $X^2(3) = 12.47$, $p = .01$, $W = .22$. Kendall's *W* provides a measure of agreement for rankings by participants and can

be interpreted as an effect size. In this case, the rankings are based on the SOA values. As the value of *W* increases, the higher degree of agreement or order can be found. A value equal to 1 indicates unanimous order by all participants, while a value of 0 indicates random ordering. Generally, the degree of agreement between the ordering of SOAs of the websites at the various accuracy thresholds was not high. Planned Wilcoxon tests indicated that social networks/blogs required significantly longer SOAs than both search and shopping websites to obtain 75% and 90% accuracy. While there was a cumulative difference in SOAs between search and news websites, this effect was not significant for any accuracy thresholds. As accuracy thresholds increased, the SOAs for news websites became significantly shorter than the SOAs for social networks/blogs, reaching significance at the 90% accuracy threshold. No differences between search and news websites or shopping and search websites were noted. Overall, medium to large effect sizes were found for the significant comparisons. See Table 9 for websites category comparisons of SOAs across accuracy thresholds and Table 10 for the results of planned Wilcoxon statistical tests.

TABLE 9

MEAN SOAS FOR EACH WEBSITE CATEGORY ACROSS ACCURACY THRESHOLDS

	QUEST Mean Estimates of SOAs (SD) in milliseconds			
Category	60%	75%	90%	Overall
News	397 (291)	641 (338)	839 (228)	626 (204)
Search	272 (235)	516 (343)	805 (200)	531 (220)
Shopping	257 (239)	449 (296)	698 (312)	468 (231)
SN/Blogs	387 (234)	772 (240)	953 (74)	704 (150)

TABLE 10

STATISTICAL TESTS COMPARING SOAS OF WEBSITES ACROSS ACCURACY THRESHOLDS

Comparison	*Z*	*p*-value	r^2
Search Websites - News Websites - 60%	-2.21	.03	.13
Search Websites - Social Networks/Blogs - 60%	-2.09	.04	.11
Shopping Websites - Social Networks/Blogs - 60%	-1.93	.05	.10
Shopping Websites - News Websites - 60%	-1.81	.07	.09
Search Websites - Shopping Websites - 60%	-.77	.45	.02
Social Networks/Blogs - News Websites - 60%	-.20	.84	<.01
Shopping Websites - Social Networks/Blogs - 75%	-3.06	< .01*	.25
Search Websites - Social Networks/Blogs - 75%	-2.90	< .01*	.22
Shopping Websites - News Websites - 75%	-1.97	.05	.10
Search Websites - News Websites - 75%	-1.49	.14	.06
Social Networks/Blogs - News Websites - 75%	-1.41	.16	.05
Search Websites - Shopping Websites - 75%	-.77	.45	.02
Shopping Websites - Social Networks/Blogs - 90%	-3.22	< .01*	.27
Search Websites - Social Networks/Blogs - 90%	-2.94	< .01*	.23
Social Networks/Blogs - News Websites - 90%	-2.58	.01*	.17
Shopping Websites - News Websites - 90%	-2.13	.03	.12
Search Websites - Shopping Websites - 90%	-1.21	.23	.04
Search Websites - News Websites - 90%	-.77	.45	.02

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

Next, website categories were compared against the upright and inverted scenes at the 60% accuracy threshold. For these analyses, Friedman's tests were conducted because of the lack of consistent normality across the website categories and accuracy thresholds.

First, social networks/blogs were compared against the scene categories. At the 60% accuracy threshold, the results indicated significant differences between the SOAs for the three categories, $X^2(2) = 17.16$, $p < .01$, $W = .45$. The Wilcoxon tests indicated social networks/blogs

required significantly longer SOAs than upright scenes or inverted scenes to achieve 60% accuracy.

At the 75% accuracy threshold, the results indicated significant differences between the SOAs for the three categories, $X^2(2) = 26.00$, $p < .01$, $W = .68$. Wilcoxon tests indicated social networks/blogs required significantly longer SOAs than upright scenes or inverted scenes to achieve 75% accuracy.

At the 90% accuracy threshold, the results indicated significant differences between the SOAs for the three categories, $X^2(2) = 30.63$, $p < .01$, $W = .81$. Wilcoxon tests indicated social networks/blogs required significantly longer SOAs than upright scenes or inverted scenes to achieve 90% accuracy. When comparing across accuracy thresholds for social networks/blogs, as the thresholds increased, the SOAs significantly increased, $p < .01$. See Table 11 for comparisons between social networks/blogs, upright scenes, and inverted scenes. See Table 12 for the results of the planned Wilcoxon tests.

Overall, as the accuracy threshold increased, the degree of agreement for the order of SOAs of upright scenes, inverted scenes, and social networks/blogs increased and indicated a high level of agreement for the order of SOAs across participants. For the individual pairwise comparisons, 5 of 6 effect sizes were large while the remaining comparison had a medium effect size (social networks/blogs – inverted scenes – 60%).

TABLE 11

MEAN SOAS FOR SCENE CATEGORIES AND THE SOCIAL NETWORKS/BLOGS CATEGORY

	QUEST Mean Estimates of SOAs (*SD*) in milliseconds			
Category	60%	75%	90%	Overall
Upright Scenes	109 (28)	248 (190)	539 (202)	299 (123)
Inverted Scenes	198 (128)	383 (244)	593 (212)	391 (170)
SN/Blogs	387 (234)	772 (240)	953 (74)	704 (150)

TABLE 12

STATISTICAL TESTS COMPARING SOCIAL NETWORKS/BLOGS, UPRIGHT, AND INVERTED SCENES

Comparison	*Z*	*p*-value	r^2
Social Networks/Blogs - Upright Scenes - 60%	-3.78	< .01*	.38
Social Networks/Blogs - Inverted Scenes - 60%	-3.02	< .01*	.24
Social Networks/Blogs - Upright Scenes - 75%	-3.82	< .01*	.38
Social Networks/Blogs - Inverted Scenes - 75%	-3.42	< .01*	.31
Social Networks/Blogs - Inverted Scenes - 90%	-3.82	< .01*	.38
Social Networks/Blogs - Upright Scenes - 90%	-3.54	< .01*	.33

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

Next, news websites were compared against the scene categories. At the 60% accuracy threshold, the results indicated significant differences between the SOAs for the three categories, $X^2(2) = 13.05$, $p < .01$, $W = .34$. Wilcoxon tests indicated news websites required significantly longer SOAs than upright scenes or inverted scenes to achieve 60% accuracy.

At the 75% accuracy threshold, the results indicated significant differences between the SOAs for the three categories, $X^2(2) = 11.68$, $p < .01$, $W = .31$. Wilcoxon tests indicated news websites required significantly longer SOAs than upright scenes or inverted scenes to achieve 75% accuracy.

At the 90% accuracy threshold, the results indicated significant differences between the SOAs for the three categories, $X^2(2) = 18.11$, $p < .01$, $W = .48$. Wilcoxon tests indicated news websites required significantly longer SOAs than upright scenes or inverted scenes to achieve 90% accuracy. When comparing across accuracy thresholds for news websites, as the thresholds increased, the SOAs significantly increased, $p < .05$. See Table 13 for comparisons between news websites, upright scenes, and inverted scenes. See Table 14 for the results of the planned Wilcoxon tests.

Overall, the degree of agreement for the order of SOAs across accuracy thresholds of upright scenes, inverted scenes, and news websites indicated a moderate level of agreement for the order of SOAs. For the individual pairwise comparisons, effect sizes were moderate to large. The effect sizes for news websites and upright scenes were larger than the effect sizes for news websites and inverted scenes.

TABLE 13

MEAN SOAS FOR SCENE CATEGORIES AND THE NEWS WEBSITE CATEGORY

	QUEST Mean Estimates of SOAs (*SD*) in milliseconds			
Category	60%	75%	90%	Overall
Upright Scenes	109 (28)	248 (190)	539 (202)	299 (123)
Inverted Scenes	198 (128)	383 (244)	593 (212)	391 (170)
News	397 (291)	641 (338)	839 (228)	626 (204)

TABLE 14

STATISTICAL TESTS COMPARING NEWS WEBSITES, UPRIGHT, AND INVERTED SCENES

Comparison	*Z*	*p*-value	r^2
News Websites - Upright Scenes - 60%	-3.50	< .01*	.32
News Websites - Inverted Scenes - 60%	-2.50	.01*	.16
News Websites - Upright Scenes - 75%	-2.94	< .01*	.23
News Websites - Inverted Scenes - 75%	-2.46	.01*	.16
News Websites - Inverted Scenes - 90%	-3.22	< .01*	.27
News Websites - Upright Scenes - 90%	-2.98	< .01*	.23

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

Next, search websites were compared against the scene categories. At the 60% accuracy threshold, the results indicated significant differences between the SOAs for the three categories, $X^2(2) = 13.05$, $p < .01$, $W = .34$. Wilcoxon tests indicated search websites required significantly longer SOAs than upright scenes, but not inverted scenes to achieve 60% accuracy.

At the 75% accuracy threshold, the results indicated significant differences between the SOAs for the three categories, $X^2(2) = 6.42$, $p = .04$, $W = .17$. Wilcoxon tests indicated search websites required significantly longer SOAs than upright scenes, but not for inverted scenes to achieve 75% accuracy.

At the 90% accuracy threshold, the results indicated significant differences between the SOAs for the three categories, $X^2(2) = 12.74$, $p < .01$, $W = .34$. Wilcoxon tests indicated search websites required significantly longer SOAs than upright scenes, and inverted scenes to achieve 90% accuracy. When comparing across accuracy thresholds for search websites, as the thresholds increased, the SOAs significantly increased, $p < .01$. See Table 15 for comparisons

between search websites, upright scenes, and inverted scenes. See Table 16 for the results of the planned Wilcoxon tests.

The degree of agreement for the order of SOAs across accuracy thresholds of upright scenes, inverted scenes, and search websites indicated a low to moderate level of agreement for the order of SOAs. For the individual pairwise comparisons, effect sizes were moderate to large when comparing search websites to upright scenes. When comparing to inverted scenes, small to moderate effect sizes were found.

TABLE 15

MEAN SOAS FOR SCENE CATEGORIES AND THE SEARCH WEBSITE CATEGORY

	QUEST Mean Estimates of SOAs (*SD*) in milliseconds			
Category	60%	75%	90%	Overall
Upright Scenes	109 (28)	248 (190)	539 (202)	299 (123)
Inverted Scenes	198 (128)	383 (244)	593 (212)	391 (170)
Search	272 (235)	516 (343)	805 (200)	531 (220)

TABLE 16

STATISTICAL TESTS COMPARING SEARCH WEBSITES, UPRIGHT, AND INVERTED SCENES

Comparison	*Z*	*p*-value	r^2
Search Websites - Upright Scenes - 60%	-3.38	< .01*	.30
Search Websites - Inverted Scenes - 60%	-1.25	.21	.04
Search Websites - Upright Scenes - 75%	-2.78	.01*	.20
Search Websites - Inverted Scenes - 75%	-1.25	.21	.04
Search Websites - Upright Scenes - 90%	-3.22	< .01*	.27
Search Websites - Inverted Scenes - 90%	-2.78	.01*	.20

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

Finally, shopping websites were compared against the scene categories. At the 60% accuracy threshold, the results indicated no significant differences between the SOAs for the three categories, $X^2(2) = 2.95$, $p = .23$, $W = .08$. No differences were noted between the categories at the 60% accuracy threshold.

At the 75% accuracy threshold, the results indicated significant differences between the SOAs for the three categories, $X^2(2) = 7.90$ $p = .02$, $W = .21$. Wilcoxon tests indicated shopping websites required significantly longer SOAs than upright scenes, but not for inverted scenes to achieve 75% accuracy.

At the 90% accuracy threshold, the results indicated no significant differences between the SOAs for the three categories, $X^2(2) = 5.16$, $p = .08$, $W = .14$. No differences were noted between the categories at the 90% accuracy threshold. When comparing across accuracy thresholds for shopping websites, as the thresholds increased, the SOAs increased significantly, $p < .05$. See Table 17 for comparisons between shopping websites, upright scenes, and inverted scenes. See Table 18 for the results of the planned Wilcoxon tests.

Generally, the degree of agreement for the order of SOAs across accuracy thresholds of upright scenes, inverted scenes, and shopping websites indicated a low level of agreement for the order of SOAs. For the individual pairwise comparisons, effect sizes were generally small when comparing shopping websites to inverted scenes. When comparing to upright scenes, small to moderate effect sizes were found.

TABLE 17

MEAN SOAS FOR SCENE CATEGORIES AND THE SHOPPING WEBSITE CATEGORY

	QUEST Mean Estimates of SOAs (*SD*) in milliseconds			
Category	60%	75%	90%	Overall
Upright Scenes	109 (28)	248 (190)	539 (202)	299 (123)
Inverted Scenes	198 (128)	383 (244)	593 (212)	391 (170)
Shopping	257 (239)	449 (296)	698 (312)	468 (231)

TABLE 18

STATISTICAL TESTS COMPARING SHOPPING WEBSITES, UPRIGHT, AND INVERTED SCENES

Comparison	*Z*	*p*-value	r^2
Shopping Websites - Upright Scenes - 60%	-2.25	.02	.13
Shopping Websites - Inverted Scenes - 60%	-1.17	.24	.04
Shopping Websites - Upright Scenes - 75%	-2.25	.02*	.13
Shopping Websites - Inverted Scenes - 75%	-.64	.52	.01
Shopping Websites - Upright Scenes - 90%	-1.61	.11	.07
Shopping Websites - Inverted Scenes - 90%	-1.41	.16	.05

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

Subjective Comments

Participants were asked if they had strategies to detect targets during their RSVP tasks. A wide variety of criteria was considered by participants. Only for shopping websites was the content of text explicitly noted and this was limited to currency (e.g., dollar sign) or prices of merchandise (*n*=4). Almost all of the participants (*n*=17) mentioned using pictures as part of their strategy for detecting shopping websites. When using pictures to detect shopping websites, participants commonly used pictures containing merchandise or models posing with products. When the layout of a shopping website was considered as part of their strategy, the

comments regarding layout typically targeted the number of pictures on a web page or the arrangement of those pictures. In a few comments, specific elements such as shopping cart details (e.g., add to cart, quantity, size) or sale signs/advertisements were noted. Only two participants associated the brightly colored web pages with shopping.

Participant strategies for detecting search websites had one notable difference from strategies for detecting other website categories. There was a distinct lack of pictures used in the task. No participants mentioned pictures are part of their strategy. Most commonly, participants looked for only a search bar (*n*=17) on a website or lists of text that resembled search results (*n*=15).

Almost all participants (*n*=18), stated they used pictures as part of their strategy for detecting news websites. They commonly used pictures containing individuals in formalwear (e.g., politicians) or some type of event (e.g., natural disasters, wars). Several participants (*n*=12) remarked they looked for text in an arrangement with pictures. Finally, only a few participants (*n*=5) described the number of pictures or the size of the pictures.

Finally, pictures were important for detecting social network/blogs. Participants used pictures that looked "amateur," were self-taken or contained people that were not famous. Overall, 13 of 19 participants used pictures of people as a detection cue. Seven participants noted that the background color was used for detection, while 3 participants specifically looked for layouts similar to Facebook, Twitter, or MySpace.

Discussion

The primary purpose of this study was to establish whether websites could be detected with above chance accuracy from information derived from stimulus presentations of a fixation

or less. In this study, the maximum presentation time of the stimuli was 140 ms, followed by an ISI, when applicable. With the use of an ISI, a 300 ms icon following the presentation of a stimulus equated approximately 100 ms of additional exposure time for perceptual processing (Loftus, et al., 1985). An icon has been assumed to last approximately 250 ms, where the quality of the icon degrades over time. Given this, SOAs exceeding 390 ms to 440 ms would have approximately 240 ms of stimulus exposure over the course of the SOA. Additional time during the SOA can be attributed to processing of information. This exposure time, regardless of the length of SOA, is within an average fixation during scene perception (Rayner, 2009).

Detection of websites

Overall, the results of this study support that information extracted from websites was sufficient to detect targets with above chance accuracy with exposures of a fixation or less. However, the mean SOA of website categories at each accuracy threshold exceeded those of upright and inverted scenes.

Mean SOAs from 250 ms to 400 ms were required to achieve 60% accuracy, while SOAs under 800 ms were sufficient to detect any of the website categories with 75% accuracy. To achieve 90% accuracy on the detection task, many participants demonstrated that SOAs equivalent to the maximum SOA allowed in this study may have not been sufficient. This suggests that in order to reach 90% accuracy, SOAs exceeding one second would be needed if participants were capable of achieving the accuracy threshold with the presentation times used in this study.

The overall website SOAs for each accuracy threshold were similar or greater than the 390-440 ms (presentation time plus icon) that would be required for the full 240 ms of

exposure to the stimulus. It was apparent that participants required the full exposure time plus additional processing to detect websites with any degree of above chance accuracy.

Gist and reduced gist standards

There has been extensive debate in the literature as to what constitutes gist. A concrete operational definition of gist has yet to be defined within scene perception literature regarding gist. Several studies agree that gist, or the semantic meaning, can be extracted from a scene within fixation (Fei-Fei, et al., 2007; Greene & Oliva, 2009a; Potter, 1975; Potter, 1976; Potter & Fox, 2009; Potter & Meng, 2008; Oliva, 2005; Oliva & Schyns, 2000; Oliva & Torralba, 2007). Wolfe, et al. (2011) labeled gist as a combination of basic attributes, basic categorization, and spatial layout. Gist is characterized by many aspects of scene perception (Oliva, 2005). It was proposed that conceptual gist (schemas or semantic information) arises from the perceptual information of the scene, or the perceptual gist, which describes the global properties of a scene. This occurs during or after the presentation time of the scene. In essence, perceptual gist enhances conceptual gist. In exploring perceptual gist, the spatial envelope theory describes several attributes of scenes which can be implicated in rapid categorization of a scene at a basic or superordinate category. For both perceptual and conceptual gist, prior research suggested that neither type of gist needs to rely on processing of objects within the scene for gist extraction.

The primary goal of this research was not to establish a concrete definition of gist. However, for the purposes of this research, a working, operational definition was necessary. A common working definition has been what is extracted within a fixation (Fei-Fei, et al., 2007; Greene & Oliva, 2009a; Potter, 1975; Potter, 1976; Potter & Fox, 2009; Potter & Meng, 2008;

Oliva, 2005). Given this, the simplest definition of gist would be the semantic information that is extracted within 266 ms (Rayner, 2009). Within a fixation, scenes can be rapidly perceived where semantic information is extracted and objects can be processed (Fei-Fei, et al., 2007; Greene & Oliva, 2009a; Liu, et al., 2009; Potter, 1975; Potter, 1976; Potter & Meng, 2008; Wolfe, et al., 2011). Previous research indicated that objects can be decoded as fast as 100 ms (Liu et al., 2009). It would be expected that at least a few objects could be decoded during a 266 ms fixation.

Another possible way to examine gist content in this experiment would be to look at the differences between the upright and inverted scene conditions at the three accuracy thresholds. Inversion of scenes has previously been shown to reduce or change gist, if not completely eliminate it (Diamond & Carey, 1986; Epstein, et al., 2005; Evans & Treisman, 2005; Kelley, et al., 2003; Meng & Potter, 2008; Shore & Klien, 2000; Torralba & Oliva, 2003). The use of inverted scenes had several advantages. Features such as spatial structure, color, and luminance remain consistent regardless of orientation and allows the cost of inversion in terms of SOA to be attributed to the reduction of the gist extracted from a scene.

The upright scene conditions required SOAs of 109 ms, 248 ms, and 539 ms to achieve 60%, 75%, and 90% detection accuracy. Overall, the SOAs for inverted scenes were significantly longer than the SOAs for upright scenes. For inverted scenes, SOAs of 198 ms, 383 ms, and 593 ms were required to achieve 60%, 75%, and 90% detection accuracy. At the 90% accuracy threshold, no differences were noted between upright and inverted scenes.

At 60% accuracy, inverted scenes required 89 ms longer for detection than for upright scenes. At 75% accuracy, inverted scenes required an additional 135 ms. However, at 90%

accuracy, the SOA for inverted scenes was greater by 53 ms. From these results, the inversion of scenes resulted in significant decreases in the gist of a scene. Because the visual quality of the scene was not modified and confounds were not introduced, it would be suspected that the inversion resulted in decreased quality of conceptual gist. This provided a relative measure of how detection performance was impeded by inverting scenes. As accuracy thresholds increased and duration increased, the relative reduction in gist decreased. At 90%, the influence of gist appeared less important for detection of upright and inverted scene targets, indicating that participants likely relied more on local feature processing.

This does not conclusively demonstrate that the inverted condition eradicated gist from the scenes. Participants were still capable of detecting inverted scene targets with 60% accuracy with SOAs less than 200 ms. Moreover, this also does not conclusively exclude that object processing occurred for upright scenes at the 60% threshold.

In this study, participants were able to detect upright scene targets with above chance accuracy during RSVP tasks. At 60% accuracy, the SOAs in this study were similar to previous research (Potter, 1975; Potter, 1976). For 75% accuracy, the SOAs in this study were slower than those found by Potter (1975; 1976), but similar to those found by Intraub (1981). This discrepancy may be explained by the richness of the cue used and also the use of interstimulus intervals (ISI). In this study, the cue used was only a simple descriptor of the basic category (beach or mountain), which was defined at the beginning of the study. Previous research has demonstrated differences in the richness of the cue for rapid detection tasks where a richer description or image lent to better detection performance (Intraub, 1981). In Potter (1975; 1976), richer, individually tailored descriptions were used to describe the targets.

Potter & Fox (2009) found presentation times of 240 ms equated to performance of 78% accuracy. When the task used SOAs with 160 ms presentation time followed by an ISI, participants still achieved 65% accuracy in detecting targets during the RSVP tasks that contained a single image per frame. The findings of the current study were consistent with previous research using RSVP tasks using ISIs.

In this study, participants were able to achieve 90% accuracy, but with longer SOAs. In previous research, detection performance did not reach 90% accuracy, where SOAs of 400 ms and 720 ms resulted in detection accuracy of 80% (Potter & Fox, 2009). In comparison with previous research and the differences between upright and inverted scene SOAs, the results of this study suggest that for the baseline upright scene condition, gist, derived from basic category descriptors, was extracted from scenes with decreasing influence on detection performance as SOAs increased.

The SOAs for upright scenes at the 60% and 75% accuracy thresholds were completed in the time course of a fixation. However, at the 90% accuracy threshold, the SOAs exceeded the maximum exposure to the stimulus, suggesting that participants relied, in part, on additional processing. Based on comparisons with previous research, the upright scene conditions appear to be representative of gist extraction and inverted scenes likely represented reduced gist scenario. This allowed direct comparisons with the website conditions to evaluate the degree of gist extraction.

Website gist

The definition of gist typically states that it is not dependent upon the processing of local objects and that it involves the extraction of semantic information from stimulus

exposures lasting no longer than a single fixation. For websites, SOAs were typically longer than for both upright and inverted scenes. Generally, these differences resulted in moderate to large effect sizes, particularly for social networks/blogs and news websites. The effect sizes for search and shopping websites were typically smaller. Participants were able to detect websites within a fixation, but they typically required additional processing time to accurately make decisions regarding the presence of website targets. Theoretically, it is possible that gist was extracted from websites given that participants had above chance detection accuracy from less than a single fixation's worth of stimulus exposure. However, it is also possible that detection was partially dependent upon local web elements. If the detection of websites was dependent upon local web elements, it appears that the type of gist extract was conceptually, not perceptually based.

The idea that gist was not sufficient for website detection without the processing of local features has exceptions. For both shopping and search websites, gist may have had a greater impact, albeit more weakly than reduced gist extracted from inverted scenes. At the 60% accuracy threshold, SOAs for shopping and search websites did not differ significantly from SOAs for inverted scenes. Moreover, when these website categories were compared with inverted scene SOAs at higher thresholds, they were shorter, though not significantly due to Holm-Bonferroni corrections. For shopping websites, the mean SOA at the 60% threshold did not differ from the SOA of the upright scene condition, either. Taken together, gist may have contributed to the detection of shopping websites, if not search websites, also.

It was possible that both of these website categories were distinctive enough from news websites and social networks/blogs that they were easier to detect. This idea would be

consistent where basic-level categorizations maximize in-group similarities and out-group dissimilarities. While they may have been distinctive, processing of a few local features or web elements may have been sufficient for detection of websites. The SOAs for both allow for the processing of local features at any accuracy threshold. The pictures on shopping websites were primarily merchandise, while most of the search website stimuli lacked pictures. In both of these cases, detection of local features, or lack thereof, may have aided detection.

Participants were queried about the strategies they deployed during the detection task. After the experiment, participants stated they used pictures of merchandise or models, large sale signs, or currency/prices to detect the presence of shopping websites. For search websites, participants commonly remarked that web pages that resembled search results (e.g., large bodies of blue text with few images) or web pages with only a search bar in the center were sufficient to detect search websites. In either situation, it is likely that processing many features would not be necessary to reach 60% accuracy given the distinctness of these website categories.

Similar scenarios may also explain the performance related to social networks/blogs and news websites. Both of these categories required SOAs that were longer than inverted and upright scenes, regardless of accuracy threshold. Participants commonly remarked that images were used to detect both types of websites. Social networks/blogs were commonly detected with profile like pictures, pictures that contained people, or pictures that looked like they were taken by an amateur. Participants stated that detection of news websites relied commonly on pictures of politicians (people), violent events (war, natural disasters), sports, or other newsworthy events. For social networks/blogs, participants commonly remarked they looked

for amateurish pictures instead of the professional pictures they expected on news websites. Prior research has found that targets and distractors that share similar global properties may make detection tasks more difficult (Greene & Oliva, 2009b). While global features were not measured in this study, it may be possible that similarities between social networks/blogs and news contributed to the longer SOAs.

The relationships between search and shopping websites with natural scenes, in combination with the specific conceptual schemas described by participants, indicated they possibly had more distinct semantic information used during the detection of websites.

With the information provided, the remaining hypotheses can be evaluated. Hypothesis H_1 stated that participants would be able to detect targets for scene and website categories with stimulus exposures not exceeding a fixation, but upright natural scenes would have lower SOAs than websites. Participants successfully detected targets for upright scene and inverted scene categories, regardless of accuracy threshold. At 60% and 75% accuracy thresholds, participants successfully detected website targets, but required SOAs possibly exceeding one second to reach 90% accuracy in this study. Additionally, accurate detection of websites needed additional processing time to achieve similar performance levels as both upright and inverted scenes. Across all types of categories in the study, there were corresponding significant increases in SOAs as the accuracy thresholds increased. However, the difficulty of detecting websites with 90% accuracy leads to only a partial rejection of the null hypothesis.

These results partially supported the rejection of the null hypothesis for hypothesis H_2, which stated that website categories with higher agreement would be more readily detected during the RSVP tasks. In the website classification study, which resulted in the web stimuli

used in this study, shopping websites had significantly higher agreement than the other categories. To fully reject the null hypothesis, the SOAs for detecting shopping websites should be significantly shorter than the other SOAs, with no significant differences between the other website SOAs across the accuracy thresholds. In this study, the SOA for detecting shopping websites was lower than those for social networks/blogs, but only at 75% and 90% accuracy thresholds. However, a similar significant relationship was found for the SOAs of search websites and social networks/blogs for 75% and 90% accuracy. In accordance with the prediction of the hypothesis, this result was unexpected. Moreover, a significant difference in SOAs was noted between news and social networks/blogs at the 90% accuracy threshold. Finally, search and shopping websites related SOAs were similar, regardless of accuracy threshold.

Hypothesis H_3 stated the SOAs associated with upright scenes will be shorter than those associates with inverted scenes, which, in turn would be similar or shorter than SOAs for website categories. The SOAs associated with upright scenes were significantly shorter than SOAs for inverted scenes at 60% and 75% accuracy thresholds. However, at the 90% accuracy threshold, the SOAs were similar between upright and inverted scenes. In regards to website SOAs, both upright and inverted scene SOAs were significantly shorter than both news and social networks/blogs SOAs regardless of accuracy threshold. Search website SOAs did not differ from inverted scene SOAs at 60% and 75% accuracy thresholds. At 90%, both upright and inverted scene SOAs were significantly shorter. Finally, shopping website SOAs were similar to SOAs for both upright and inverted scene categories with the exception of the 75% accuracy

level, where upright scene SOA was significantly shorter than the shopping website SOA. These findings supported the partial rejection of the null hypothesis.

Study Limitations

There were limitations to this study. Due to a configuration file error, fewer social network/blog websites were available during the RSVP tasks. As a result, Facebook website stimuli were excluded from sampling. Additionally, this increased the number of times the distractors were viewed since this study randomly selected distractors with replacement. This may or may not have been beneficial for participants. Due to the higher number of presentations, it may have been possible to disregard social networks/blogs during trials when they were distractors, thus decreasing the difficulty of those trials. Moreover, the exclusion of Facebook may have increased the SOAs, but the probability of a single Facebook stimulus being selected as target was .36%. Having fewer possible targets available for random selection without replacement may have been beneficial. The impact of this was examined from data collected in the subsequent study, which found that accurate detection of social networks/blogs was significantly less than 75% accuracy when all of the stimuli were used, but no differences were noted across the conditions when social networks/blogs were compared to the other three categories. No differences were noted in the percentage of agreement for social network/blog stimuli calculated in the classification study. See Appendix H for more information.

Due to the duration of the study, an inverted website condition was not included. Inclusion of an inverted website condition would allow analogous comparisons between scenes and websites, instead of needing to infer the relationship from the SOAs of upright and inverted

scenes Future research should explore the effect of inverting websites to determine whether this would lead to similar increases in SOAs.

Conclusions

This study demonstrated that participants were capable of detecting websites prompted by basic categories with above chance accuracy from stimulus exposures of 240 ms or less. In order to do so, it required significantly longer SOAs than for upright and inverted natural scene detection. This suggests that additional processing time, and in some cases additional stimulus exposure, was needed to facilitate the detection.

Study 3 – Effects of Size and Resolution on the Perception of Websites

Study 2 showed that participants were capable of detecting websites displayed for less than a fixation, but required additional processing time to differentiate between the categories. Overall, participants exhibited the best performance for detecting shopping websites in comparison with the other websites. Participant comments indicated they may have relied on web elements such as pictures to detect the presence of websites during the RSVP tasks.

The influence of local features can possibly be moderated by manipulating the size and resolution of the websites. Torralba (2009) found that participants were able to elicit gist and identify a number of objects when viewing scenes with detail limited by the number of available pixels within the stimulus. While performance decreased as the number of pixels in the image decreased, yet the size of the image was maintained, they were still capable of detecting targets that lacked fine features. Oliva and Schyns (2000) found that participants could recognize scene types based on coarse color blobs derived from the original scenes.

Previous research has noted that small resolution screens may be particularly problematic for displaying websites (Chittaro, 2006). Users have limited viewing of websites and must scroll and zoom to see the entire website in detail. Other research has shown that creating thumbnails of websites to display on small device screens have benefits where the website layouts were preserved, but participants commonly commented about poorly rendered text and indistinguishable images (Lam & Baudisch, 2005). Moreover, they found that participants liked having access to the original layouts, but layout by itself was not adequate to find desired content in some cases.

While these are not website related studies, they do note differences in screen size and image quality of video. Screen size has been implicated in learning studies where the amount of information learned during a video based learning exercise was diminished by smaller screen sizes (1.65”) versus larger screen sizes: 2.28” and 3.78” (Maniar, Bennett, Hand, & Allan, 2008). Other research has shown that image size and resolution affect the acceptability of the image when being viewed in a video feed (Knoche, McCarthy, & Sasse, 2005). Content displayed in 120 x 90 px and high quality typically was found to be unacceptable by viewers if viewed with poor bandwidth. Common complaints included not being able to discern detail or read text.

The purpose of this experiment was to determine the influences of the local features, such as pictures, and how size and resolution influences detection of websites, which is applicable to how websites are commonly displayed on these different types and sizes of displays. From previous literature and the findings of the previous studies, several hypotheses can be formed.

H_1: Participants' ability to discriminate between categories of websites will be moderated by size and resolution. Participants' ability to discriminate between categories will decrease as size decreases. Additionally, their ability to discriminate will decrease as resolution decreases. When resolution can be downsampled in terms of image size, the effects of resolution should be greater in smaller stimulus sizes than larger sizes. Due to this, it is anticipated that an interaction between size and resolution will exist, but the nature of the interaction is not predicted.

H_2: Discriminability at smaller sizes will be better for website categories with higher agreement. Based on the classification study, shopping should be more readily discriminated from other types of websites.

H_3: Discriminability at lower resolutions will be better for website categories with higher agreement. Based on the classification study, shopping should be more readily discriminated from other types of websites.

H_4: Participants will exhibit either a tendency to respond that targets were present less or more often toward specific website categories at smaller sizes and lower resolutions. This hypothesis is exploratory. A conservative bias indicates that more information is needed to respond more often that the target is present. It is anticipated that a bias will be exhibited and will differ when size and resolution change.

Method

Participants

Twenty-nine participants, 18 years of age and older, were recruited from Wichita State University or the Wichita Metropolitan Area. Participants volunteered for the study through

SONA Systems, the University experiment management system, or by contacting the researcher directly. Participants volunteering via SONA Systems were compensated with course credit. Out of the 29 participants, three participants were screened due to suboptimal near visual acuity, two were screened due to suboptimal color vision, two were removed due to significantly higher nonresponse rates, and one participant was excluded for not following instructions provided for the experiment. Overall, the remaining 21 participants, (5 males, 16 females; *M Age* = 22.00 years, *SD* = 3.33 years) were included in the analyses. See Appendix I for more details regarding the screening of participants. Please see Appendix J for more participant demographics.

Materials

Equipment. The same experimental setup was used in this study as in Study 2. There were changes to the RSVP software for size and resolution conditions.

Questionnaire. The same questionnaire from Study 2 was used to gather information regarding demographics, computer usage, web usage, and video gaming experience.

Stimuli. The same four website categories were used in this study as in Study 2. For each type of website, 276 screenshots were used. Web stimuli were displayed in three sizes. The Large size was 1,024 x 772 px, the Medium size was 512 x 386 px, and the Small size was 128 x 96 px. The ratio of height and width equivalent to a scaled version of websites displayed in a web browser at 1,280 x 1,024 px with the vertical scrollbar and browser UI removed. At 60cm, the Large stimulus subtended 27.01 degrees horizontally and 20.52 degrees vertically, the Medium stimulus subtended 13.69 degrees horizontally and 10.34 degrees vertically, and the

Small stimulus subtended 3.44 degrees horizontally and 2.58 degrees vertically. See Figure 9 for sample stimuli at the three stimulus sizes.

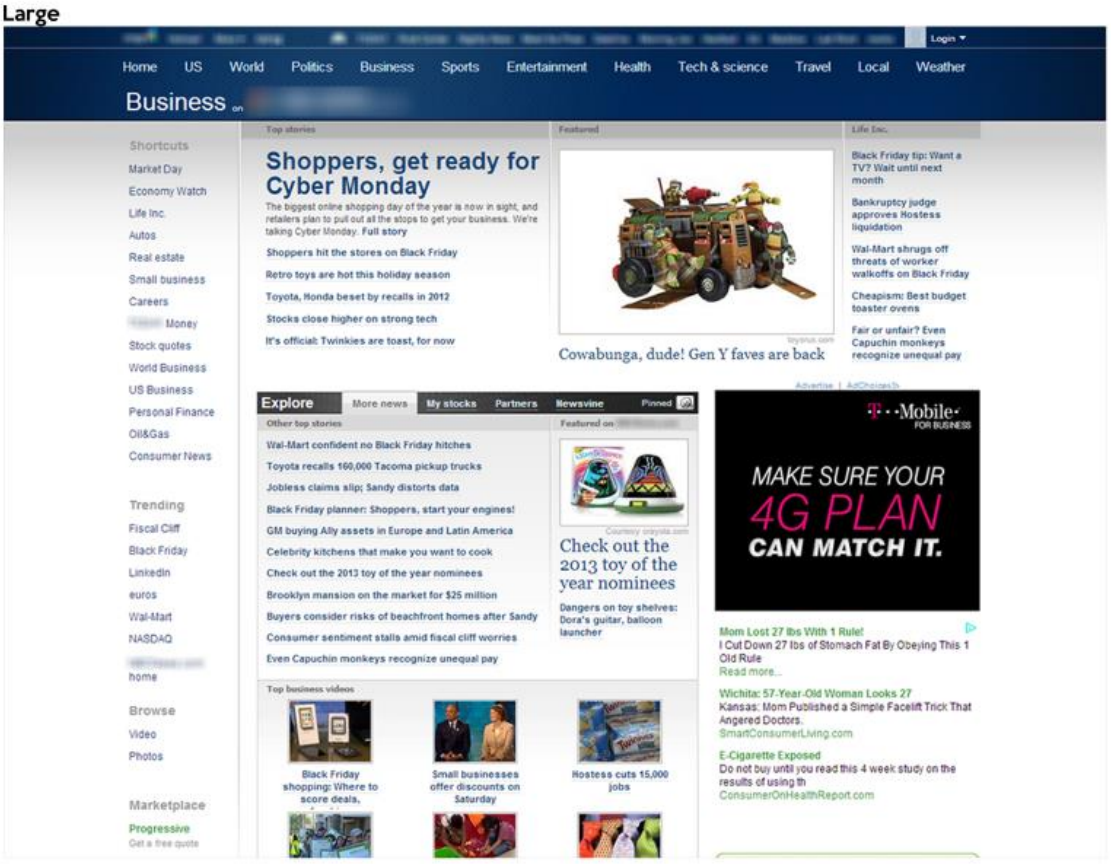

Figure 9. Relative size comparisons of the stimuli.

Procedure

Generally, the same study procedure was used in this study as in Study 2. A similar RSVP paradigm was used, except only websites were used, and size and resolution were manipulated. See Figure 10 for the schematic of the RSVP trial.

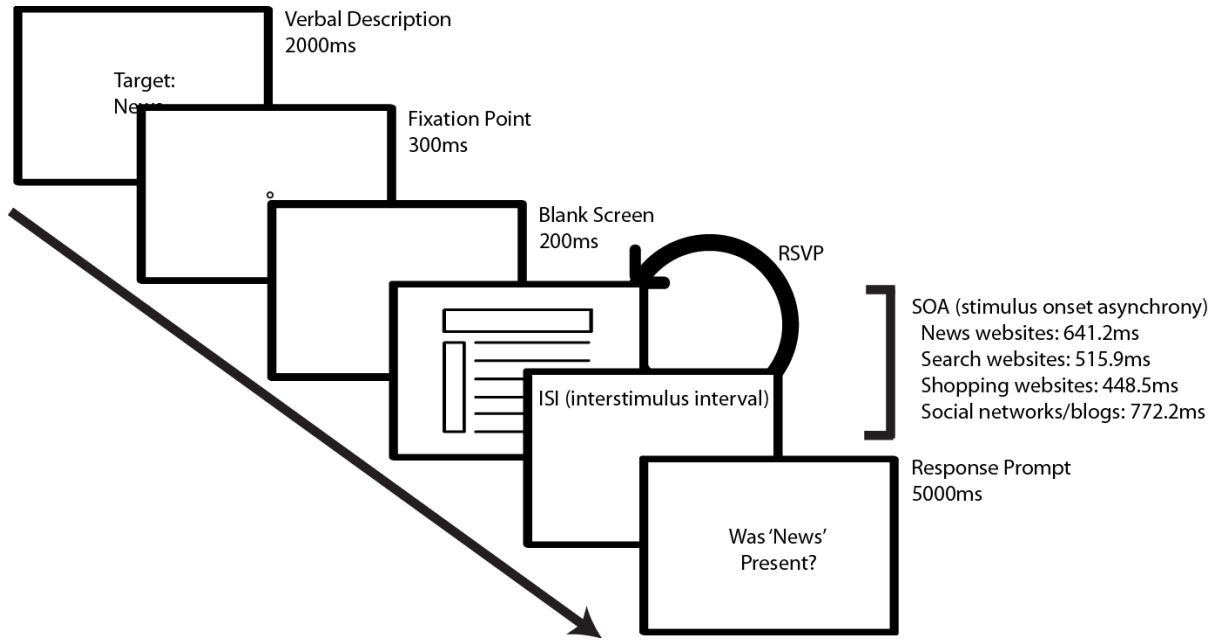

Figure 10. RSVP trial schematic for study 3.

The SOA for each category was determined in the previous study. The SOAs of the websites at the 75% accuracy threshold were used for this study. For news websites, the SOA used was 641.2 ms. The SOA used for search websites was 515.9 ms. The SOA used for shopping websites was 448.5 ms. Finally, the SOA used for social networks/blogs was 772.2 ms.

Screen display can be examined in two different dimensions: size and resolution. Since the physical size and resolution of displays used in computing today vary, various sizes of stimuli will be used in this experiment. The stimuli will be displayed in three sizes. The size conditions include Small: 128 px; Medium: 512 px; and Large: 1,024 px.

To help control the amount of visual information being conveyed, the resolutions were manipulated as well. The stimuli were displayed in three different resolutions. These resolutions include 1:1 (High), 1:4 (Moderate), and 1:8 (Low) ratios. A ratio of 1:8 is considered to be lower resolution than 1:1 or 1:4 ratios. For example, at a 1:1 ratio, a stimulus 128 px in width will contain 128 distinct pixels of data for width. At a 1:4 ratio, the Small stimulus in width would have 32 px of visual data. The amount of visual data in the stimulus was controlled using a procedure adapted from Torralba (2009). As Torralba stated, the use of Gaussian blur does not control the amount of visual information contained in the images. Subsampling maintains a better upper bound of the information contained within the stimuli. These ratios were referred to as High resolution, Moderate resolution, and Low resolution for the remainder of the study.

Smaller images were generated from the original stimuli, which were 1,250 x 942 px in size. All smaller images were 96 dpi and contained 8 bits per color channel. The smaller images were saved as png format, which uses lossless compression. The sizes of the smaller images were 16 x 12 px, 32 x 24 px, 64 x 48 px, 128 x 96 px, 256 x 193 px, 512 x 386 px, and 1,024 x 772 px. Next, each image size was upsampled to one of the three targeted stimulus sizes, depending on the resolution ratio. For the Small condition, 16 px and 32 px sizes were upsampled to the Small stimulus size for Low and Moderate resolutions. For the Medium condition, 64 px and 128 px sizes were upsampled for Low and Moderate resolutions. Finally, for the Large condition, 128 px and 256 px sizes were upsampled for the Low and Moderate resolutions. See Figures 11 and 12 for relative comparisons of size and resolution.

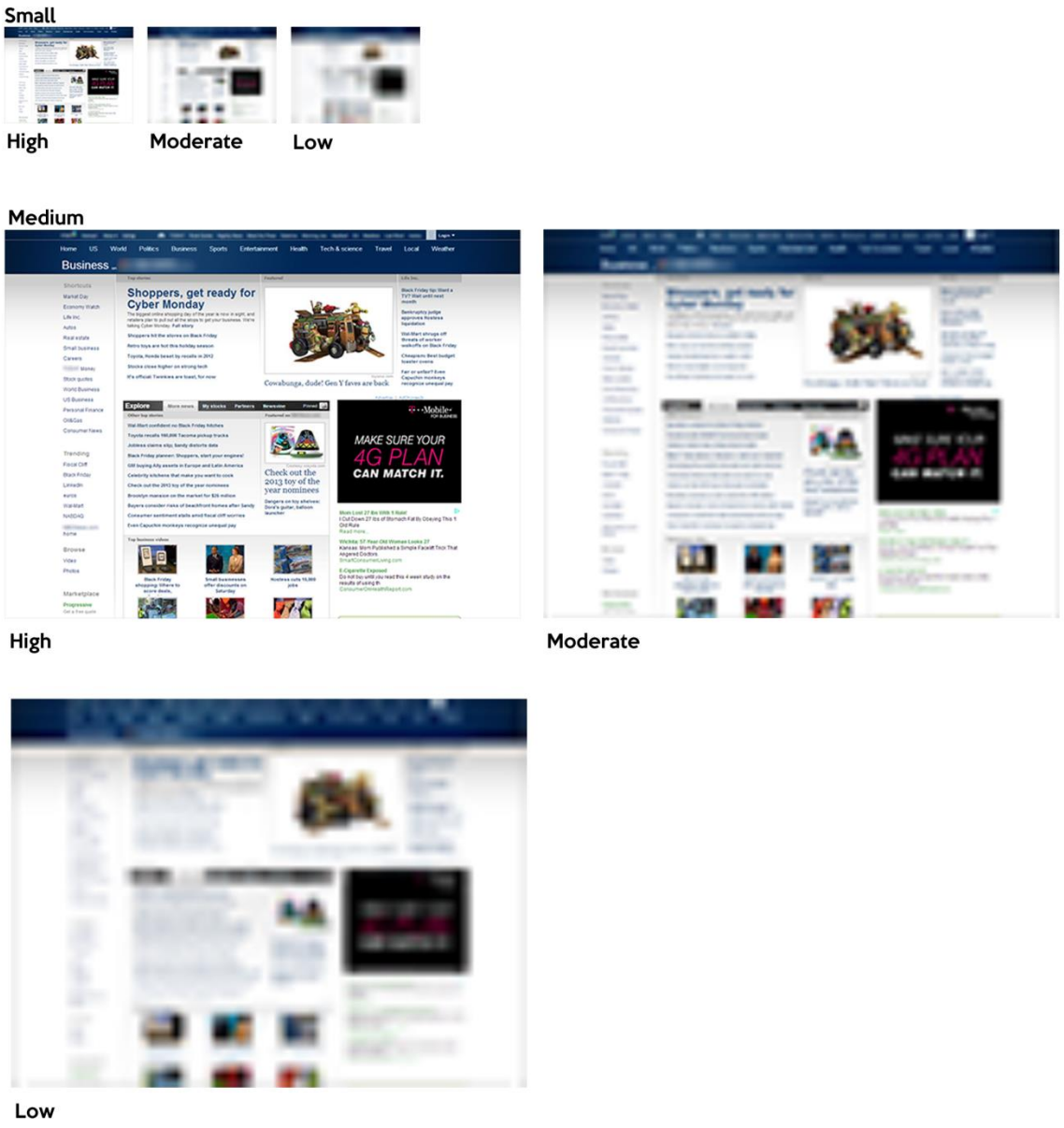

Figure 11. Comparison of size and resolution of Small and Medium images.

Figure 12. Comparison of size and resolution of Large images.

The trials were in blocks based on the stimuli size presented. For each size condition block, participants completed 16 practice trials. In each practice trial, stimuli resolutions were randomized across target-present and target-absent trials. Then, for each resolution/stimulus combination, participants completed 16 experimental trials, equaling 192 trials. Overall, participants completed 48 practice trials, which were followed by 576 experimental trials distributed equally across the three size conditions. With practice, participants performed 312 target-present and 312 target-absent trials. In each block, trials for different resolutions and whether targets were was randomized.

The entire experiment was completed over two sessions to maintain that a single session could be completed in a 2 hour period. The block order was counter-balanced over both experimental sessions to eliminate order effects. The time to complete each size block was approximately 45 minutes. Participants were instructed to take 1-2 minute breaks after every 15 minutes of RSVP trials. For the first session, one block was completed. The other two remaining blocks were completed on during the second experimental session. An extended break was provided between the two size blocks during the second experimental session.

Similar to Study 2, participants were queried about the strategies they deployed to detect target categories, completed the post-experiment questionnaire, and then were debriefed about the experiment.

Dependent Measures

Because discriminability should vary across size/resolution combinations, signal detection theory was used in this study to evaluate participants' capacity for target detection. Raw Yes/No responses were converted to hit and false alarm rates for each category. Both hit

rates and false alarm rates from response data was used to determine discriminability and bias. For this analysis, A' (A prime) was calculated for discriminability. B''_D was calculated for bias instead of d' and C.

Both A' (Equation 1), and B''_D (Equation 2) are non-parametric measures and do not carry the same priori assumptions as d' and C (Donaldson 1992; MacMillan & Creelman, 2005; Stanislaw & Todorov, 1999). Stanislaw & Todorov (1999) noted that assumptions for d' cannot be analyzed for Yes/No tasks, which were the basis of participant responses in this study. Donaldson (1992) recommended B''_D instead of B" as a measure of bias because B" may perform poorly in circumstances of low discrimination. Please see Appendix K for a secondary analysis of the hit and false alarm data with B". A' values for A' ranged from 0 to 1. Values of .5 indicate that targets could be discerned from nontargets and 1 indicated flawless discriminability. Values for B''_D ranged from -1 to 1 where 0 indicated no bias, -1 indicated liberal bias where participants more often answered "yes" and 1 indicated conservative bias where participants answered "no" more often.

$$A' = .5 + \left[sign(H - F) \frac{(H - F)^2 + |H - F|}{4\max(H, F) - 4HF} \right] \quad (1)$$

$$B''_D = \frac{(1 - H)(1 - F) - HF}{(1 - H)(1 - F) + HF} \quad (2)$$

A small correction was applied to the overall accuracy, hit rate, and false alarm rate (MacMillan & Creelman, 2005; Stanislaw & Todorov, 1999). The correction adjusted rates equal to 100% by subtracting 1/2n, where n is number of samples in the proportion. Likewise, to adjust rates equal to 0%, 1/2n was added. The corrections were performed to alleviate issues with calculating A' and B''_D that occur when hit and false alarm rates were equal to 0 and/or 1.

A prime, hit rates, false alarm rates, and accuracy rates are bounded measures. Bounded measures such as proportions have been noted to not meet the assumptions of parametric statistics, which require continuous data. Logit transformations have been found to perform better than arcsine transformations when transforming proportion data and also results in more interpretable comparisons (Warton & Hui, 2011). To unbound these measures, logit transformations can be applied. The results were reported in the raw and backtransformed data. The bias measure for B''_D was not transformed to stabilize variances due to the possible range of values.

Results

The data was checked for normality and parametric statistical assumptions. See Appendix G for more information. Supplementary results have been provided in Appendix K for web usage, video game usage, and an additional measure of bias, B".

Sensitivity

For each participant, sensitivity was measured with A'. Because A' resembles a proportion and was bounded (0, 1), logit transformed values were used for analysis to meet parametric assumptions. Normality of the data was checked across each combination of website type (news, search, shopping, social networks/blogs), stimulus size (Small, Medium, Large), and stimulus resolution (High, Moderate, Low). Overall, each resolution combination with the Large search website category, and the Small size, Low resolution condition for the shopping website category, exhibited negative skewness, $p < .01$. However, due to robustness and sufficient degrees of freedom, a repeated-measures ANOVA was conducted (DeCarlo,

1997; Fidell & Tibachnick, 2004). See Tables 19 and 20 for means and standard deviations of A'.

See Table 5 for means and standard deviations of logit A'.

TABLE 19

MEANS FOR SENSITIVITY (A')

Size	Resolution	News Websites	Search Websites	Shopping Websites	Social Networks/Blogs
Small	High	.59 (.20)	.81 (.14)	.82 (.11)	.71 (.16)
Small	Moderate	.59 (.19)	.77 (.15)	.64 (.21)	.57 (.17)
Small	Low	.55 (.15)	.76 (.13)	.60 (.17)	.57 (.20)
Medium	High	.79 (.17)	.78 (.18)	.81 (.16)	.73 (.16)
Medium	Moderate	.70 (.21)	.80 (.17)	.84 (.11)	.74 (.18)
Medium	Low	.74 (.14)	.80 (.15)	.79 (.12)	.66 (.18)
Large	High	.83 (.11)	.83 (.21)	.87 (.11)	.79 (.14)
Large	Moderate	.76 (.15)	.84 (.16)	.82 (.13)	.75 (.12)
Large	Low	.75 (.17)	.81 (.18)	.83 (.11)	.74 (.17)

TABLE 20

BACKTRANSFORMED MEANS FOR SENSITIVITY (LOGIT A')

Size	Resolution	News Websites	Search Websites	Shopping Websites	Social Networks/Blogs
Small	High	.61 [.53, .69]	.85 [.81, .88]	.84 [.80, .87]	.74 [.66, .80]
Small	Moderate	.60 [.50, .70]	.81 [.75, .85]	.67 [.57, .75]	.58 [.48, .67]
Small	Low	.55 [.46, .64]	.78 [.72, .83]	.60 [.50, .69]	.58 [.47, .69]
Medium	High	.82 [.76, .88]	.83 [.77, .88]	.85 [.79, .89]	.76 [.71, .81]
Medium	Moderate	.74 [.66, .80]	.83 [.79, .87]	.87 [.83, .90]	.78 [.71, .83]
Medium	Low	.76 [.70, .81]	.84 [.79, .87]	.81 [.76, .85]	.68 [.60, .75]
Large	High	.86 [.81, .90]	.89 [.83, .93]	.89 [.86, .93]	.82 [.77, .86]
Large	Moderate	.79 [.74, .83]	.88 [.84, .91]	.85 [.82, .88]	.76 [.72, .81]
Large	Low	.78 [.73, .83]	.85 [.80, .88]	.86 [.82, .89]	.77 [.71, .82]

Note: CIs are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008).

A 3-way repeated measures ANOVA was conducted to examine the differences between the four website categories at the three stimulus sizes and three resolutions for A'. Mauchly's tests of sphericity were conducted for the three main effects and the four interactions. Sphericity was noted for the three-way interaction between stimulus type, stimulus size, and stimulus resolution, $W(77) = .01$, $p = .02$. A Greenhouse-Geisser adjustment was conducted to compensate for the presence of sphericity.

The results of the repeated-measures ANOVA indicated significant main effects for stimulus type, $F(3,60) = 20.93$, $p < .01$, partial $\eta^2 = .51$, stimulus size, $F(2,40) = 42.29$, $p < .01$, partial $\eta^2 = .68$, and stimulus resolution, $F(2,40) = 22.30$, $p < .01$, partial $\eta^2 = .53$. The results indicated significant interactions also for stimulus type and size, $F(6,120) = 4.12$, $p < .01$, partial $\eta^2 = .17$, and stimulus size and resolution, $F(4,80) = 2.50$, $p = .05$, partial $\eta^2 = .05$. Interactions for stimulus type and resolution, $F(6,120) = 1.27$, $p = .23$, partial $\eta^2 = .06$, and the three-way interaction for stimulus type, stimulus size, and stimulus resolution were not significant, $F(6,120) = 1.82$, $p = .10$, partial $\eta^2 = .08$. For the main effects, large effect sizes were found. For the significant interactions, medium and large effect sizes were found.

First, the significant interaction for stimulus type and stimulus size was further evaluated with planned comparisons to determine which conditions significantly differed. In the Small size condition, participants demonstrated greater sensitivity for detecting search websites than news websites, shopping websites, and social networks/blogs. As the size of the stimulus increased to the Medium size, participants' sensitivity was greater for shopping websites than both news websites and social networks/blogs. In the Large size, participants were more sensitive for detecting shopping and search websites than social networks/blogs.

Additionally, participants demonstrated higher sensitivity for detecting shopping websites over news websites in the Large size condition. When comparing differences in participant sensitivity for stimulus size within each stimulus type, participants demonstrated greater sensitivity for detecting news websites, shopping websites, and social networks/blogs in the Medium or Large size conditions than the Small size condition. For the significant comparisons, large effect sizes were found for each. See Figure 13 for the interaction between stimulus size and type. See Tables 21 and 22 for planned pairwise comparisons between stimulus types for the different sized stimuli. Essentially, sensitivity decreased significantly when news, shopping, and social networks/blogs stimuli were displayed in the Small size, but not the two larger sizes.

In the Small size condition, no differences were noted between news and shopping websites in comparison with social networks/blogs. In the Medium size, no differences were noted in sensitivity for search websites compared with news websites, shopping websites, or social networks/blogs. In the Large size, no differences were noted between search websites compared with news or shopping websites. No differences were noted between news websites and social networks/blogs. Finally, no significant differences were noted in sensitivity for search website stimuli, regardless of stimulus size. Overall, for the significant comparisons, large effect sizes were found.

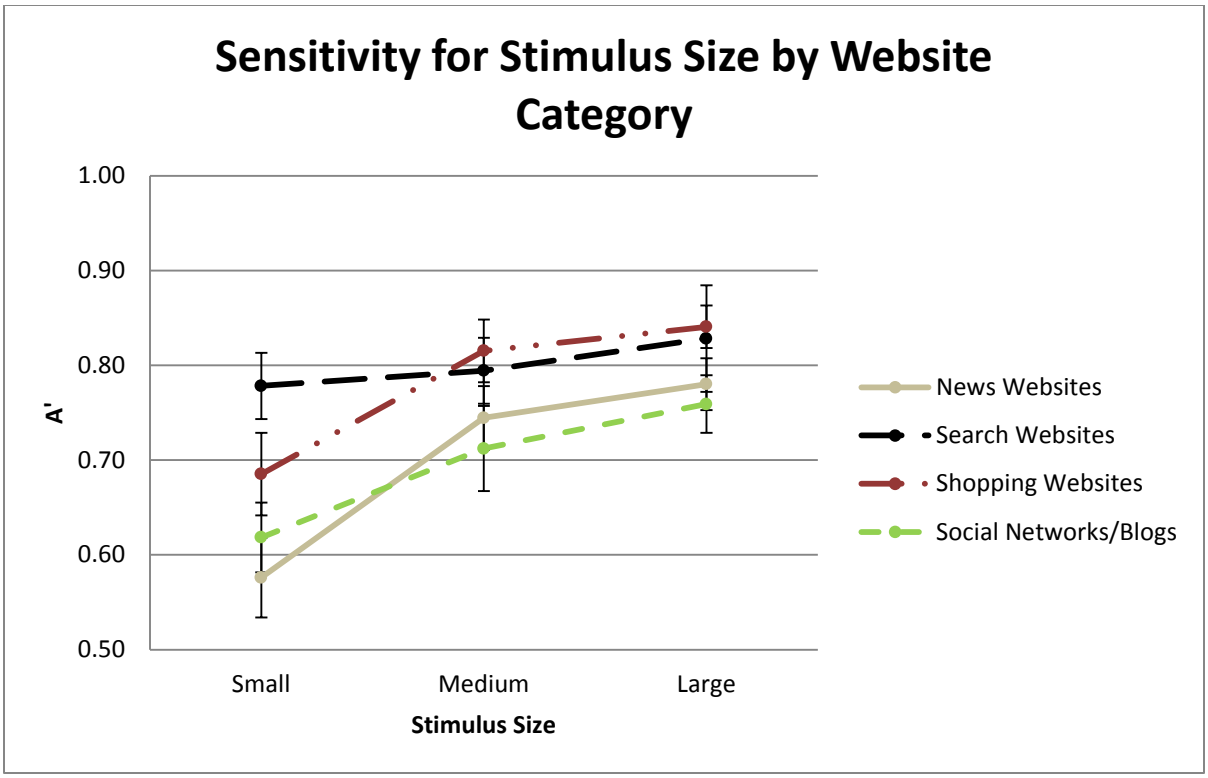

Figure 13. Sensitivity for detecting websites in the various stimulus sizes.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

TABLE 21

STATISTICAL TESTS COMPARING SENSITIVITY FOR STIMULUS SIZES AND WEBSITE STIMULI

Comparison	*t*	*p*-value	r^2	MD
News Website: Small - Large	-8.41	< .01*	.78	-.22
News Website: Small - Medium	-7.28	< .01*	.73	-.19
Shopping Website: Small - Large	-6.15	< .01*	.65	-.15
Social Networks/Blogs: Small - Large	-6.07	< .01*	.65	-.15
Shopping Website: Small - Medium	-5.94	< .01*	.64	-.13
Social Networks/Blogs: Small - Medium	-3.21	< .01*	.34	-.11
Search Website: Small - Large	-2.53	.02	.24	-.06
Social Networks/Blogs: Medium - Large	-1.97	.06	.16	-.04
Search Website: Medium - Large	-1.89	.07	.15	-.04
News Website: Medium - Large	-1.72	.10	.13	-.04
Shopping Website: Medium - Large	-1.58	.13	.11	-.03
Search Website: Small - Medium	-1.08	.29	.06	-.02

Note: * indicates a significant comparison. Mean differences are differences in backtransformed means. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

TABLE 22

STATISTICAL TESTS COMPARING SENSITIVITY FOR WEBSITES AT DIFFERENT STIMULUS SIZES

Comparison	*t*	*p*-value	r^2	MD
Small: News Website - Search Website	-6.80	< .01*	.70	-.22
Small: Search Website - Social Networks/Blogs	6.09	< .01*	.65	.18
Large: Shopping Website - Social Networks/Blogs	4.90	< .01*	.54	.09
Small: News Website - Shopping Website	-4.30	< .01*	.48	-.13
Medium: Shopping Website - Social Networks/Blogs	3.95	< .01*	.44	.10
Small: Search Website - Shopping Website	3.78	< .01*	.42	.10
Large: Search Website - Social Networks/Blogs	3.75	< .01*	.41	.09
Large: News Website - Shopping Website	-3.25	< .01*	.35	-.06
Medium: News Website - Shopping Website	-3.17	.01*	.33	-.07
Medium: Search Website - Social Networks/Blogs	2.98	.01	.31	.09
Small: Shopping Website - Social Networks/Blogs	2.59	.02	.25	.08
Large: News Website - Search Website	-2.53	.02	.24	-.06
Medium: News Website - Search Website	-2.40	.03	.22	-.06
Medium: News Website - Social Networks/Blogs	1.99	.06	.16	.03
Small: News Website - Social Networks/Blogs	-1.98	.06	.16	-.05
Large: News Website - Social Networks/Blogs	1.51	.15	.10	.03
Medium: Search Website - Shopping Website	-.45	.66	.01	-.01
Large: Search Website - Shopping Website	.04	.97	< .01	< .01

Note: * indicates a significant comparison. Mean differences are differences in backtransformed means. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

Next, the significant interaction for stimulus size and stimulus resolution was further evaluated with planned comparisons to determine which conditions significantly differed. Participants exhibited higher sensitivity for detecting targets in the high resolution when displayed in both the Small and Large sizes. However, no differences were noted in sensitivity when the resolution was degraded. In the Medium size conditions, sensitivity was similar,

regardless of stimulus resolution. See Table 5 for pairwise comparisons between stimulus sizes and stimulus resolutions.

Overall, sensitivity was lower in the Small size condition than the Large size condition, regardless of resolution. At both Low and Moderate resolutions, sensitivity was lower in the Small size condition than for the Medium size condition. Between the Medium and Large sizes, no significant differences were noted in sensitivity. Generally, large effect sizes were found for the significant comparisons. See Figure 14 for the interaction between stimulus size and resolution. See Tables 23 and 24 for pairwise comparisons between stimulus sizes and stimulus resolutions.

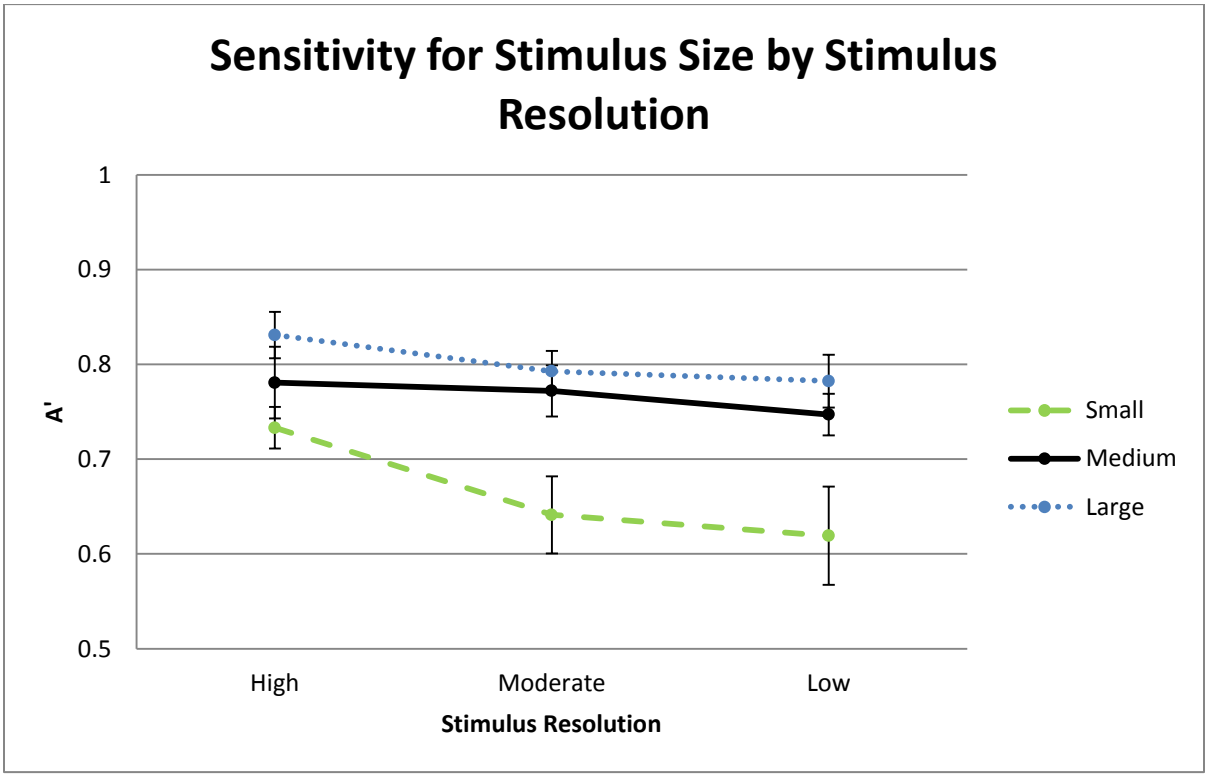

Figure 14. Sensitivity across the different stimulus size and resolutions.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

TABLE 23

STATISTICAL TESTS COMPARING SENSITIVITY FOR RESOLUTIONS AT DIFFERENT STIMULUS SIZES

Comparison	*t*	*p*-value	r^2	MD
Small: High - Low	4.69	< .01*	.52	.14
Small: High - Moderate	3.76	< .01*	.41	.10
Large: High - Low	3.46	< .01*	.37	.05
Large: High - Moderate	3.38	< .01*	.36	.04
Medium: High - Low	2.41	.03	.23	.04
Medium: Moderate - Low	2.19	.04	.19	.03
Small: Moderate - Low	1.12	.28	.06	.04
Large: Moderate - Low	.93	.36	.04	.01
Medium: High - Moderate	.43	.67	.01	.01

Note: * indicates a significant comparison. Mean differences are differences in backtransformed means. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

TABLE 24

STATISTICAL TESTS COMPARING SENSITIVITY FOR SIZES AT DIFFERENT STIMULUS RESOLUTIONS

Comparison	*t*	*p*-value	r^2	MD
Moderate: Small - Large	-6.18	< .01*	.66	-.15
High: Small - Large	-5.85	< .01*	.63	-.10
Low: Small - Large	-5.52	< .01*	.60	-.18
Moderate: Small - Medium	-5.15	< .01*	.57	-.14
Low: Small - Medium	-4.96	< .01*	.55	-.14
High: Medium - Large	-2.66	.02	.26	-.05
High: Small - Medium	-2.49	.02	.24	-.05
Low: Medium - Large	-2.36	.03	.22	-.04
Moderate: Medium - Large	-1.18	.25	.07	-.02

Note: * indicates a significant comparison. Mean differences are differences in backtransformed means. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

The conditions for Small size, High resolution, Medium size, Moderate resolution, and Large size, Low resolution used the same source images (128 px), thus they contained a similar

amount of visual information prior to being rendered in their final sizes of Small, Medium, and Large. A repeated measures ANOVA was conducted to determine if differences existed in the logit of A' for these three conditions. The results of the ANOVA indicated significant differences between the three conditions, $F(6,120) = 6.28$, $p < .01$, partial $\eta^2 = .24$. The effect size was large.

Planned comparisons showed that for the Medium and the Large size conditions, no differences were found, $p < .05$. However, sensitivity was higher for Large sizes than the Small size (*MD* = .05, $p < .01$). Similarly, sensitivity was higher for the Medium than the Small size (*MD* = .04, $p < .01$). This suggests that even though the resolution was higher for the Small size, the increased size of the stimuli sufficiently mitigated decreases in resolution found in the larger sized stimuli.

The sensitivity of size and resolution conditions for each of the stimulus types were compared with the logit of .5, which indicated whether the targeted categories could not be distinguished from the distractors in the RSVP tasks. The results of one-sample *t*-tests indicated that participants could not distinguish targets from distractors in the Small size, Low and Moderate resolutions for news websites, shopping websites, and social networks/blogs. Moreover, the same result was found for Small size, High resolution news website condition. At Medium and Large sizes, regardless of resolution, participants had sensitivity indicating they were capable of distinguishing targets from the distractors. Finally, the results indicated that for search, participants had significantly higher sensitivity, regardless of size or resolution. For each of the significant comparisons, a large effect size was found. See Figure 15 for the comparisons across all 36 conditions. See Appendix L, Table 55 for planned pairwise comparisons of A'.

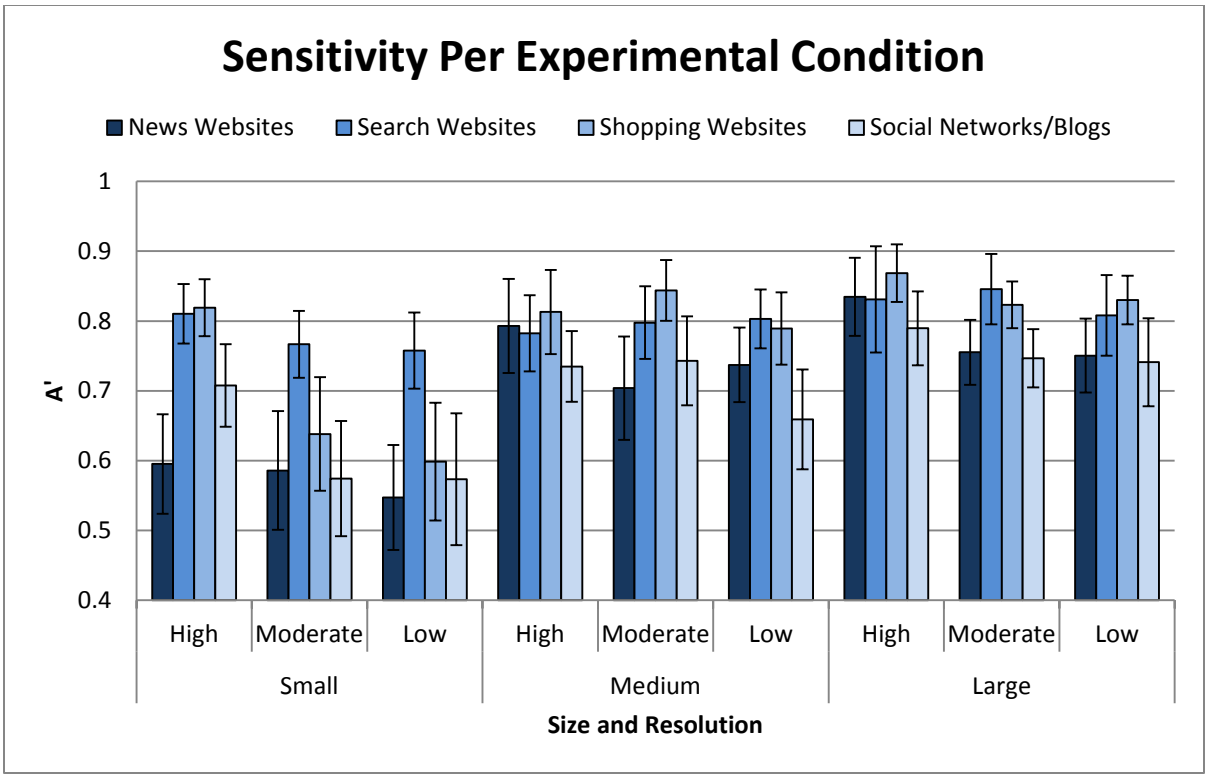

Figure 15. Sensitivity for detecting websites in each of the different size and resolutions.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

Bias

For each participant, bias was measured with B''_D. For the bias measure B''_D, normality of the data was assessed across each combination of website type (news, search, shopping, social networks/blogs), stimulus size (Small, Medium, Large), and stimulus resolution (High, Moderate, Low). Generally, most distributions exhibited some degree of negative skewness. Overall, all distributions of the resolution conditions for the largest sized social networks/blogs stimuli exhibited significant negative skewness and leptokurtosis, $p < .01$. Similarly, the distribution for the Medium sized stimulus condition at the highest resolution was significantly negatively skewed and leptokurtic, $p < .01$. For the news website category, the Medium Moderate degraded condition distributions exhibited significant negative skewness, $p < .01$. In

the Small size for shopping websites in the Low resolution, the distribution was significantly positively skewed. See Table 25 for the means and standard deviations of B''_D.

TABLE 25

MEAN AND STANDARD DEVIATIONS FOR B''_D

Size	Resolution	News Websites	Search Websites	Shopping Websites	Social Networks/Blogs
Small	High	-.16 (.61)	.18 (.52)	.25 (.53)	.35 (.64)
Small	Moderate	-.51 (.56)	.18 (.57)	-.29 (.59)	.13 (.71)
Small	Low	.08 (.64)	.28 (.53)	-.44 (.68)	-.09 (.68)
Medium	High	.11 (.54)	.30 (.51)	.49 (.41)	.51 (.46)
Medium	Moderate	.31 (.52)	.30 (.52)	.18 (.53)	.53 (.39)
Medium	Low	.44 (.53)	.26 (.53)	.22 (.51)	.44 (.47)
Large	High	.18 (.60)	.13 (.59)	.30 (.40)	.53 (.45)
Large	Moderate	.46 (.45)	.20 (.58)	.38 (.46)	.67 (.45)
Large	Low	.39 (.53)	.35 (.48)	.28 (.51)	.55 (.50)

A 3-way repeated measures ANOVA was conducted to examine the differences between the four website categories in the three stimulus sizes and three resolutions for the bias measure, B''_D. The results of the repeated-measures ANOVA indicated significant main effects for stimulus size, $F(2,40) = 15.49$, $p < .01$, partial $\eta^2 = .44$. The main effects for stimulus type $F(3,60) = 2.67$, $p = .06$, partial $\eta^2 = .12$, and stimulus resolution, $F(2,40) = .62$, $p = .55$, partial $\eta^2 = .03$, were not significant. The results indicated significant interactions also for stimulus type and size, $F(6,120) = 30.06$, $p = .01$, partial $\eta^2 = .13$, stimulus type and resolution, $F(6,120) = 7.39$, $p < .01$, partial $\eta^2 = .27$, and stimulus size and resolution, $F(4,80) = 7.37$, $p < .01$, partial $\eta^2 = .27$. The three-way interaction for stimulus type, stimulus size, and stimulus resolution was not significant, $F(12,240) = 1.43$, $p = .15$, partial $\eta^2 = .07$. For the significant main effect, a large effect size was found. For both significant interactions contending with resolution, large effect

sizes were found. For the interaction between stimulus size and type, a moderate effect size was found. Planned comparisons were conducted on the three 2-way significant interactions.

First, the significant interaction for stimulus type and stimulus size was further evaluated. When detecting shopping website targets, participants exhibited a more conservative bias as the stimulus size increased from Small to Medium or Large. Similar results were found for social networks/blogs. When detecting news website targets, participants had significantly more conservative criterion for the Large size than the Small size, but for the Small size, participants were neutral. No differences were found in bias for search websites as stimulus size increased. When comparing across website types, participants exhibited a more conservative bias when detecting search website targets than shopping website targets in the Small size. Generally, the significant comparisons had large effect sizes. See Figure 16 for the interaction between stimulus size and type. See Tables 26 and 27 for planned pairwise comparisons of B''_D.

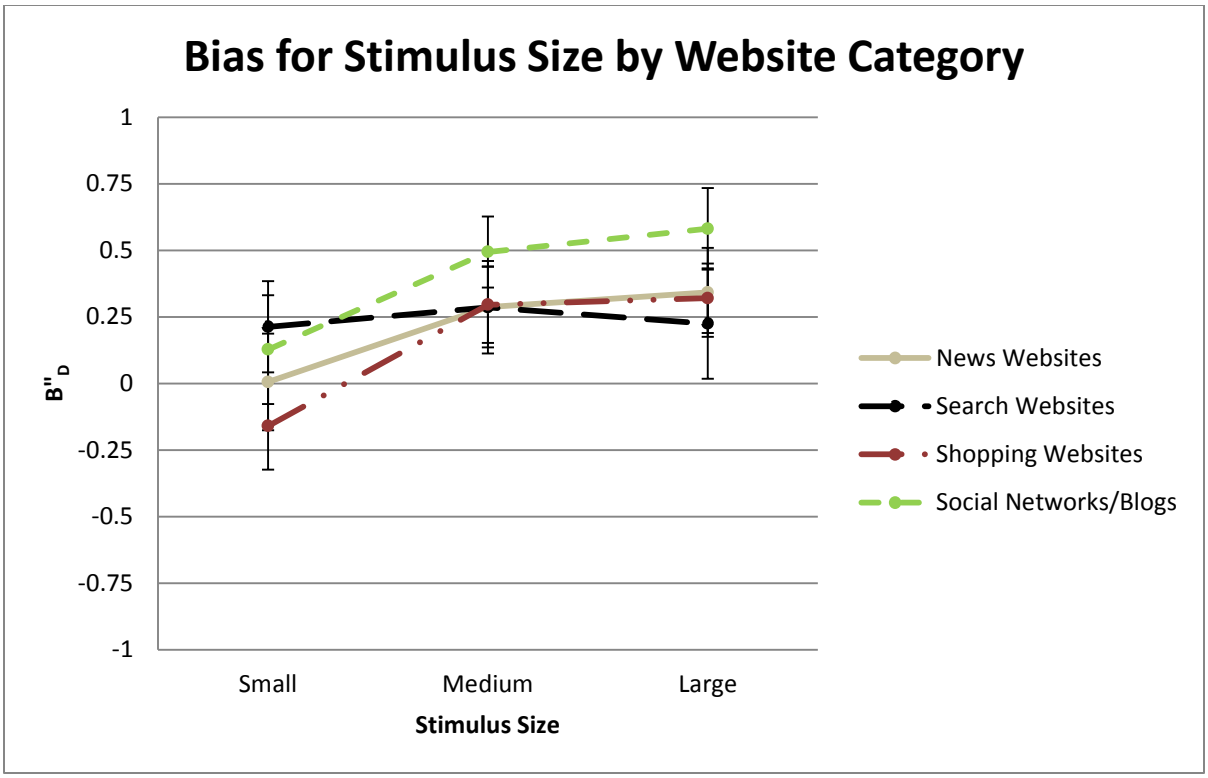

Figure 16. Bias for detecting websites across the three size conditions.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

TABLE 26

STATISTICAL TESTS COMPARING BIAS B''_D FOR STIMULUS SIZES AT DIFFERENT STIMULUS TYPES

Comparison	*t*	*p*-value	r^2	MD
Shopping Websites: Small - Large	-5.09	< .01*	.57	-.48
Shopping Websites: Small - Medium	-4.66	< .01*	.52	-.46
Social Networks/Blogs: Small - Medium	-4.08	< .01*	.45	-.37
Social Networks/Blogs: Small - Large	-3.59	< .01*	.39	-.45
News Websites: Small - Large	-3.34	< .01*	.36	-.34
News Websites: Small - Medium	-2.87	.01	.29	-.28
Social Networks/Blogs: Medium - Large	-.94	.36	.04	-.09
News Websites: Medium - Large	-.73	.46	.03	-.06
Search Websites: Small - Medium	-.61	.55	.02	-.07
Search Websites: Medium - Large	.53	.60	.01	.06
Shopping Websites: Medium - Large	-.26	.80	< .01	-.02
Search Websites: Small - Large	-.11	.91	< .01	-.01

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

TABLE 27

STATISTICAL TESTS COMPARING BIAS B''_D FOR STIMULUS TYPES AT DIFFERENT STIMULUS SIZES

Comparison	*t*	*p*-value	r^2	MD
Small: Search Websites - Shopping Websites	3.56	< .01*	.39	.37
Large: Shopping Websites - Social Networks/Blogs	-2.66	.02	.26	-.26
Large: Search Websites - Social Networks/Blogs	-2.60	.02	.25	-.36
Small: Shopping Websites - Social Networks/Blogs	-2.26	.04	.20	-.29
Medium: Shopping Websites - Social Networks/Blogs	-2.18	.04	.19	-.20
Large: News Websites - Social Networks/Blogs	-1.95	.07	.16	-.24
Medium: Search Websites - Social Networks/Blogs	-1.91	.07	.15	-.21
Medium: News Websites - Social Networks/Blogs	-1.89	.07	.15	-.21
Small: News Websites - Search Websites	-1.63	.12	.12	-.21
Small: News Websites - Shopping Websites	1.31	.21	.08	.17
Large: News Websites - Search Websites	.88	.39	.04	.12
Small: News Websites - Social Networks/Blogs	-.84	.41	.03	-.12
Large: Search Websites - Shopping Websites	-.76	.46	.03	-.10
Small: Search Websites - Social Networks/Blogs	.66	.52	.02	.09
Large: News Websites - Shopping Websites	.22	.83	< .01	.02
Medium: News Websites - Shopping Websites	-.09	.93	< .01	-.01
Medium: Search Websites - Shopping Websites	-.09	.93	< .01	-.01
Medium: News Websites - Search Websites	.01	.99	< .01	< .01

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

Next, the significant interaction for stimulus type and stimulus resolution was further evaluated. Participants exhibited differences in bias based on resolution where participants were more conservative when detecting shopping targets in the High resolution than Moderate and Low resolutions. Moreover, when examining across website types, participants exhibited a more conservative bias when detecting social networks/blogs than shopping websites when the stimuli were displayed in the Moderate resolution condition. No other differences were noted.

For the significant comparisons, large effect sizes were found. See Figure 17 for the interaction between stimulus resolution and type. See Tables 28 and 29 for planned pairwise comparisons of B''_D.

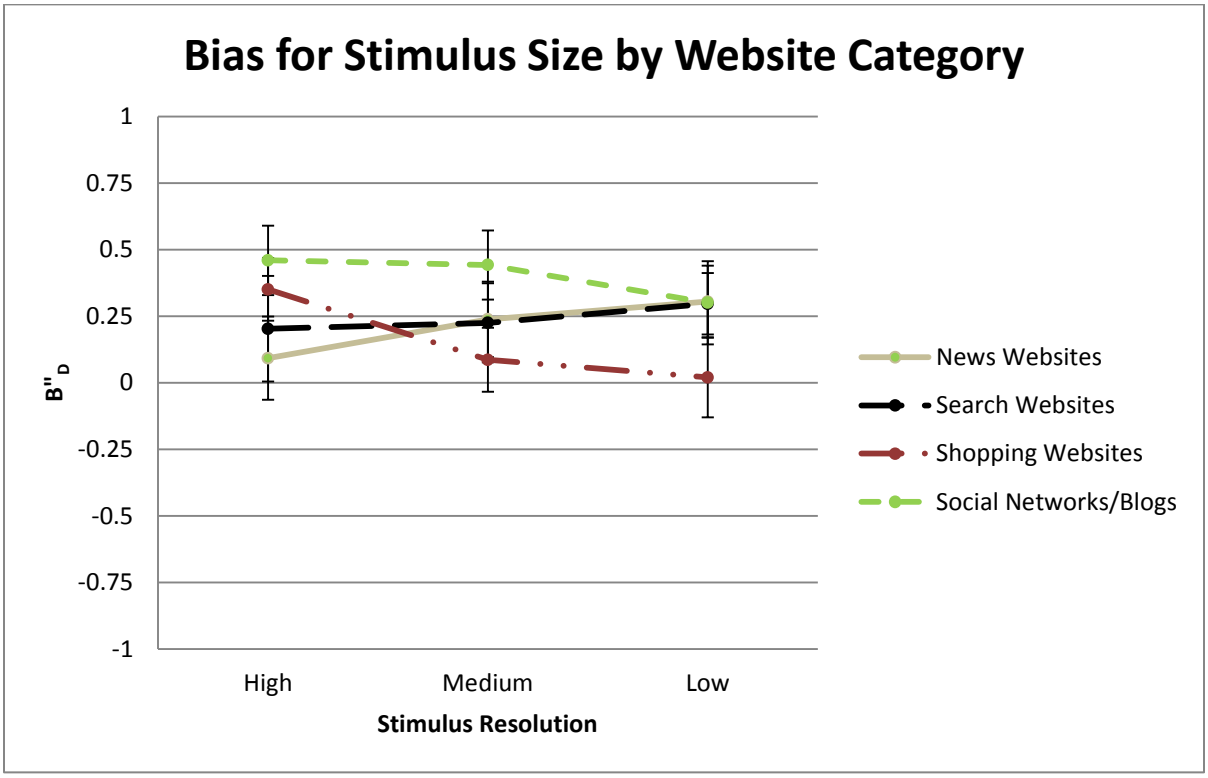

Figure 17. Bias for detecting websites across the three stimulus resolutions.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

TABLE 28

STATISTICAL TESTS COMPARING BIAS B''_D FOR STIMULUS RESOLUTIONS FOR STIMULUS TYPES

Comparison	*t*	*p*-value	r^2	MD
Shopping Websites: High - Moderate	4.21	< .01*	.47	.26
Shopping Websites: High - Low	3.58	< .01*	.39	.33
News Websites: High - Low	-2.95	.01	.30	-.21
News Websites: High - Moderate	-2.47	.02	.23	-.15
News Websites: High - Low	2.16	.04	.19	.16
Social Networks/Blogs: Moderate - Low	1.81	.09	.14	.14
News Websites: Moderate - Low	-1.59	.13	.11	-.07
Search Websites: High - Low	-1.04	.31	.05	-.09
Search Websites: Moderate - Low	-.96	.35	.04	-.07
Shopping Websites: Moderate - Low	.94	.36	.04	.07
Search Websites: High - Moderate	-.31	.76	< .01	-.02
Social Networks/Blogs: High - Moderate	.30	.77	< .01	.02

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

TABLE 29

STATISTICAL TESTS COMPARING BIAS B''_D FOR STIMULUS TYPES AT STIMULUS RESOLUTIONS

Comparison	*t*	*p*-value	r^2	MD
Moderate : Shopping Websites - Social Networks/Blogs	-4.43	< .01*	.50	-.36
High : News Websites - Social Networks/Blogs	-2.96	.01	.30	-.37
Low : Search Websites - Shopping Websites	2.81	.01	.28	.28
Low : News Websites - Shopping Websites	2.72	.01	.27	.29
High : News Websites - Shopping Websites	-2.38	.03	.22	-.26
Low : Shopping Websites - Social Networks/Blogs	-2.32	.03	.21	-.28
High : Search Websites - Social Networks/Blogs	-2.26	.04	.20	-.256
Moderate : News Websites - Social Networks/Blogs	-1.93	.07	.16	-.20
Moderate : Search Websites - Social Networks/Blogs	-1.90	.07	.15	-.22
High : Search Websites - Shopping Websites	-1.56	.13	.11	-.15
Moderate : News Websites - Shopping Websites	1.47	.16	.10	.15
High : Shopping Websites - Social Networks/Blogs	-1.29	.21	.08	-.11
Moderate : Search Websites - Shopping Websites	1.28	.22	.08	.14
High : News Websites - Search Websites	-.75	.46	.03	-.11
Moderate : News Websites - Search Websites	.12	.91	< .01	.01
Low : News Websites - Search Websites	.08	.94	< .01	.01
Low : News Websites - Social Networks/Blogs	.05	.96	< .01	< .01
Low : Search Websites - Social Networks/Blogs	-.04	.97	< .01	< -.01

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

Finally, the significant interaction for stimulus size and stimulus resolution was further evaluated. In the Low resolution condition, participants exhibited a significantly stronger conservative bias when viewing the Medium and Large sized stimuli than the Small sized stimuli. A significantly stronger bias was found when viewing the Medium and Large sized stimuli in the Moderate resolution than the Small sized stimuli. Only for the Large size did bias increase as resolution dropped from High to Moderate. For the Small and Medium sizes, bias

did not significantly differ based on resolution. Overall, large effect sizes were found for the significant comparisons. See Figure 18 for the interaction between stimulus size and type. See Tables 30 and 31 for planned pairwise comparisons of B''_D.

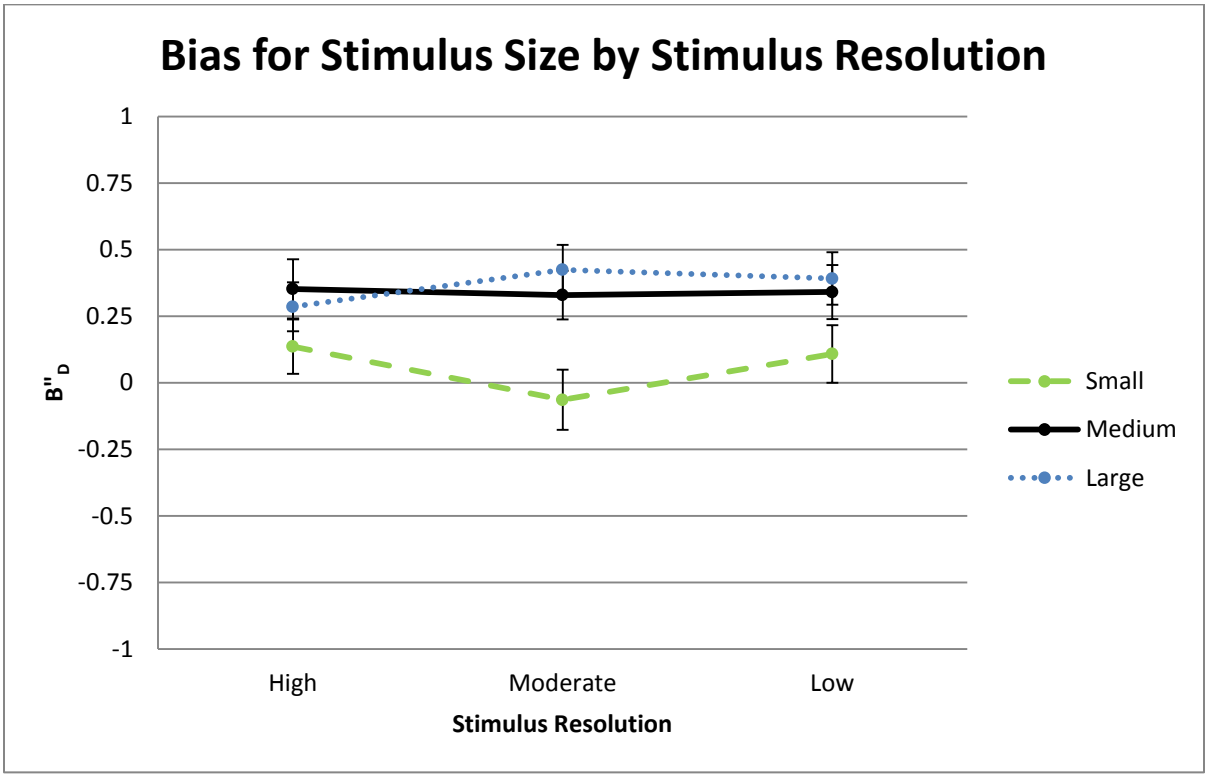

Figure 18. Bias across the size and resolution conditions.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

TABLE 30

STATISTICAL TESTS COMPARING BIAS B"$_D$ FOR STIMULUS RESOLUTIONS AT STIMULUS SIZES

Comparison	*t*	*p*-value	r^2	MD
Large: High - Moderate	-3.47	< .01*	.38	-.16
Small: High - Low	2.98	.01	.31	.23
Small: High - Moderate	2.61	.02	.26	.17
Large: Moderate - Low	1.87	.08	.15	.10
Large: High - Low	-1.39	.18	.09	-.07
Small: Moderate - Low	.74	.47	.03	.05
Medium: High - Low	.62	.55	.02	.05
Medium: High - Moderate	.59	.56	.02	.04
Medium: Moderate - Low	.14	.89	< .01	.01

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

TABLE 31

STATISTICAL TESTS COMPARING BIAS B"$_D$ FOR STIMULUS SIZES AT STIMULUS RESOLUTIONS

Comparison	*t*	*p*-value	r^2	MD
Moderate : Small - Large	-5.63	< .01*	.61	-.45
Low : Small - Large	-5.57	< .01*	.61	-.40
Low : Small - Medium	-4.52	< .01*	.51	-.35
Moderate : Small - Medium	-3.47	< .01*	.38	-.30
Moderate : Medium - Large	-1.89	.07	.15	-.15
High : Small - Medium	-1.71	.10	.13	-.16
High : Small - Large	-1.33	.20	.08	-.11
Low : Medium - Large	-1.12	.28	.06	-.06
High : Medium - Large	.60	.55	.02	.05

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

One-sample *t*-tests were conducted for the 36 conditions to compare bias against neutrality. In regards to individual conditions, 10 of 36 conditions were significant with the

Holm-Bonferroni adjustment applied. Participants had a significant conservative bias regardless of resolution, when detecting social networks/blogs at Medium and Large sizes. Additionally, participants had a significant conservative bias in their responses when detecting news websites in the Medium size, Low resolution and Large size, Moderate resolution conditions. Finally, when detecting shopping websites, participants exhibited a significantly conservative bias in the Medium size, High resolution and Large size, Moderate resolution conditions. While not significant, it was interesting that for detecting shopping websites, participants exhibited a more liberal bias at lower resolutions of the Small size. For all of the significant comparisons, large effect sizes were found. See Figure 19 for bias across the 36 conditions. See Appendix L, Table 56 for planned pairwise comparisons of B''_D.

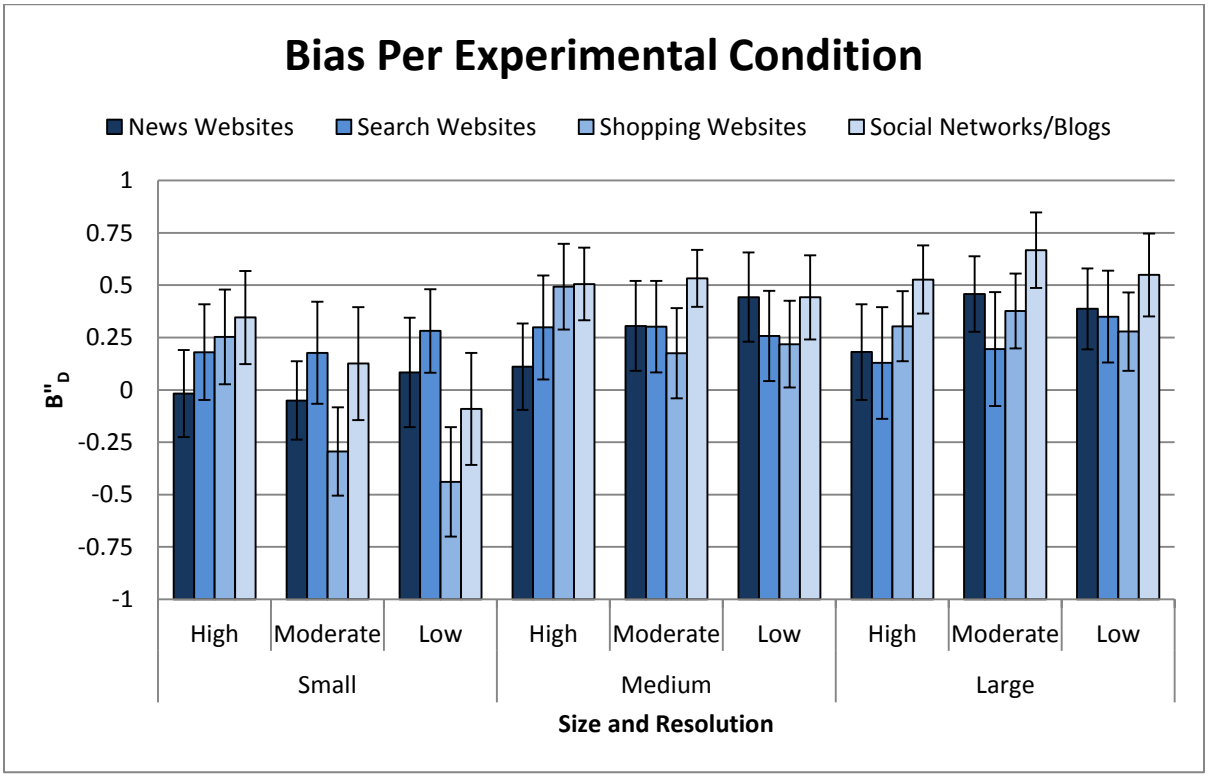

Figure 19. Bias for detecting targets each of the size and resolution conditions.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

Subjective Comments

Participants had a variety of strategies for detecting types of websites. For shopping, all of the participants had pictures as an integral part of their strategy. Specifically, most participants looked for products (*n*=17) or models (*n*=5). Other important factors included multiple items on the same page with some specificity for how the page was organized and shopping-related elements like sale advertisements, descriptions, or pricing.

Similar to shopping, participants placed an emphasis for the pictures being present in news websites (*n*=19). Of these, pictures of people (e.g., politicians) or various types of pictures like current events or sports were common comments. The majority of participants looked for text that resembled articles, while only seven looked for headlines. A select few stated the background color was important.

To detect search websites, participants focused on a few distinct aspects. These included search boxes on a single page with a white background, with a background image, or on a page with a list of links (*n*=17). Almost all participants expected research results in which the majority expected blue text. Only two participants looked for regions of advertisements.

Finally, detection of social networks/blogs was driven by the presence of profile pictures, pictures other than profile pictures, and text resembling comments. Of the 21 participants, four participants looked for either Twitter or Facebook. Finally, a few participants attempted to distinguish social networks through the usage of background color.

Discussion

Overall, participants demonstrated they were able to detect websites across almost all experimental conditions. In the experimental conditions that contained the least amount of

visual information (e.g., Small size, Low and Moderate resolutions), participants were unable to discern targets for three of the four types of websites (i.e., news websites, shopping websites, social networks/blogs) from distractors. However, even with these poor resolutions, participants were capable of distinguishing search websites from the other website categories. In the larger sizes, participants exhibited the ability to detect targets with exposures to the stimulus not exceeding a fixation.

The results of this study demonstrated that stimulus size affected participants' ability to discriminate between the four website types, but the effect was influenced by the website category. In the Small size, participants were more sensitive in detecting search websites than any other stimulus type. As the size of the stimulus increased, sensitivity for detecting search websites did not increase significantly. However, for the other three types of stimuli, significant increases in sensitivity were found from Small to Medium sizes, but not increasing the size from Medium to Large. This suggests that the information available to participants in the Medium condition was just as visually informative as the information in the Large condition.

Participants' ability to detect search websites in the small and Low resolution conditions indicated that search websites may have distinguishing properties. Search websites were unique where most search stimuli lacked pictures. Additionally, on a search home page, the page typically consisted of a search bar, surrounded by white space. On the search results page, the color of the list of hyperlinks was typically blue. These web elements possibly contributed to the higher sensitivity that was found for search websites in the smallest stimulus size. In the subjective comments for the strategies adopted by participants, they indicated white space, blue text, advertising, and centered search bars were important. No participants mentioned the

presence of pictures as part of their strategy for detecting search websites, either. In the Small size, these features would still be distinct. In the lower resolutions of the Small size, it seems probable that such features, or lack of pictures, helped distinguish search from the other types of websites. The advantage of these unique attributes was mitigated by larger stimulus sizes, where sensitivity was similar between search and shopping websites.

In the lower resolution conditions of the Small stimulus size, participants could not distinguish shopping, news, or social networks/blogs. In these conditions, the quality of web elements suffered. Given this, detection of these three types of websites appears to be dependent on web elements, such as pictures, headers, or other web elements. Generally, participants found pictures to be useful for detection of websites. In the comments about the strategies adopted by participants, they indicated that images were an important discriminatory factor for discerning news websites, shopping websites, and social networks/blogs. While this is not to discount other local features suggested by participants like headers, or combinations of text and images, the presence of pictures and the picture content appeared to be the most consistent factor mentioned in participant comments.

If merchandise pictures operated as a discriminatory factor that was not shared by the other three categories, it follows that if the content of the pictures was able to be recognized, it would aid in the detection of the website. In essence, the presence of products on the page likely acted as a single factor for participants. Evidence supported this notion in the Medium and Large conditions where sensitivity for shopping websites was significantly greater than both news and social networks/blogs and was equivalent to that of search websites. Images on both news and social networks/blogs were likely not distinct enough from each other to act as a

single discriminatory factor. For both news and social networks/blogs, the content of the pictures typically had shared attributes. For instance, both shared people, but the context and quality of such pictures differed. For social networks/blogs, profile pictures were important as were pictures of nonfamous people, while pictures of politicians and other leaders were important for news. In the Medium and Large sizes, participants had similar sensitivity for detecting social networks/blogs and news websites.

It was plausible that distinctiveness of profile pictures found on social networks/blogs or the presence of Facebook or Twitter led to similar results to shopping in the smallest stimulus size. Either of these factors may have been sufficient to increase the sensitivity in the Small size condition. If the presence of the Facebook or Twitter layout led to increased sensitivity, participants should still be able to detect targets from distractors in the lowest resolutions in the Small size condition. This was not the case. These results suggested that the increased sensitivity overall in the Small size condition was likely due to the presence of profile pictures.

The uniqueness of elements found in search websites and the images in shopping websites likely led to increased sensitivity when compared with both news websites and social networks/blogs. Essentially, the more distinct factors that participants could use resulted in higher sensitivity at all resolutions.

With High resolution stimuli, sensitivity across the three sizes was only significantly different between the smallest and largest stimuli. This indicated that even in the Small size, enough visual information was present to achieve performance similar as with the Medium sized stimuli. When resolution of the stimuli decreased, the largest decreases in sensitivity were seen in the Small condition, which indicated a significant diminishment in the amount of visual

information being utilized. This is in comparison with the Medium and Large sizes, where no significant decrease was found when resolution decreases.

If the number of pixels in the source images in the study were used as an approximation of the amount of visual data available to the participants, the amount of visual information needed to be able to discriminate between targets and distractors was surprisingly small. The Medium size, Low resolution condition used 64 px source images. This equated to 3,072 px of data. When displayed in the Medium size, it was sufficient for participants to achieve sensitivity similar to the Large size, Low resolution condition, which contained 12,288 px in the Small source image. Only when source images were 32 px and 16 px in size was there a large drop in sensitivity.

Oliva and Schyns (2000) found that color blobs provided sufficient context for participants to recognize various natural scenes. In other research, 32 x 32 px color images displayed at 256 x 256 px has been shown to be enough visual information to achieve 80% accuracy in the recognition of scenes (Torralba, 2009). When displaying 8 x 8 px color images in the 256 px size, accuracy was slightly less than 40%, which was still above chance. Only with 4 x 4 px grayscale images did participants' performance equate chance. In this study, visual information from a 32 x 24 px displayed in the Small size was insufficient to aid participants in detecting website targets from distractors, except for search websites. While there were differences in the methodological paradigms of the tasks (RSVP vs. recognition tasks), these results show that the amount of visual information needed was enlightening. These results suggested that contextual cues provided by the websites, except for search sites, provided insufficient information for increases in sensitivity.

The relationship between size and resolution was not linear. When using the same source images at different sizes and resolutions, the benefit of image resolution can be mitigated by image size. Participants were more sensitive to targets in the larger stimulus sizes than the smallest stimulus size, even though the resolutions were degraded by a factor of 4 and 8 in the larger sizes when the original source images were the same. This suggests that increasing the size of the local features may have made them more distinguishable, even though they were more degraded in resolution.

A conservative bias indicated the need for more evidence in order to respond with "yes" during a trial (Wickens, 2001). Conversely, this corresponds with an increased need for more information to respond that the target was present (Johnson, & Proctor, 2004). In this study, participants were not provided feedback regarding their performance, so changes in criterion cannot be attributed to feedback or payoff. Loschky and Larson (2010) suggested that a conservative bias indicates more variability from the distractor categories than the target category.

Across the conditions, there were differences in response bias, particularly in the Small size conditions. Participants exhibited a conservative bias when detecting targets displayed in the Medium and Large stimuli sizes, but the bias was less prominent for the Small stimuli size. For each stimulus type, except for search websites, bias generally became more conservative as the stimulus size increased. For news websites, shopping websites, and social networks/blogs, bias generally became more conservative from Small to Large stimulus sizes. As resolution decreased, the conservative bias found for shopping websites diminished.

While the bias was conservative for search websites, it did not change across the conditions. The distinct lack of change for search websites implied that the criterion used to distinguish search likely was sufficient to not significantly shift the threshold for detection either more liberally or conservatively as stimulus size and/or stimulus resolution changed. As a result, the information used by participants for detection was sufficient to maintain a similar criterion and did not to be adjusted.

The increase in conservative bias for news websites, shopping websites, and social networks/blogs as stimulus size increased implies a shift in the internal threshold used. As more visual information became available and local features were more discernible, participants shifted their criterion to a more conservative position. A similar pattern was noted for shopping when resolution decreased. With the increase in visual information, participants may have adjusted their internal response criterion in conjunction with the quality of the visual information being presented. As a result of the increased visual quality, participants appeared to have scrutinized the information more than when the quality was lower.

Overall, participants responded with a conservative bias to the larger stimulus sizes and were sensitive to detect targets from distractors, which were based on very brief exposures to the website stimuli. This posits that in the brief exposures, participants did not have enough information to respond more liberally (e.g., they required more information), but the information provided was sufficient for detection. In accordance with Loschky and Larson's (2010) reasoning, this suggests that the variability of the underlying distributions were more similar for the distractors in the smaller and lower resolution conditions, with the exception of the search websites.

The results of the study support rejecting the null hypothesis for H_1 where differences were found in sensitivity due to the conditions of the study. In respect to hypothesis H_1, participants' were capable of discriminating between the different types of websites within the SOAs used in the experiment as highlighted by participants' sensitivity. For news websites, shopping websites, and social networks/blogs, participants' sensitivity varied by stimulus size, where decreases in size resulted in decreases in sensitivity. Moreover, an interaction between stimulus size and resolution was found where the stimulus was small and resolution was not optimal, sensitivity decreased.

Hypotheses H_2 and H_3 posited that participants would have higher sensitivity for more specifically categorized websites in smaller sizes and lower resolutions. These null hypotheses should be accepted as shopping websites did not differ from news in smaller sizes or had sensitivity as high as search websites. In terms of resolution, the sensitivity for the websites generally varied together by resolution as there were no interactions contending with specific types of websites.

Finally, in evaluating the hypothesis H_4, the null hypothesis should be rejected. The differences in bias were found for shopping websites, which decreased as resolution decreased. Moreover, in smaller sizes, participants exhibited less conservative bias.

Because users were able to detect websites from categorical prompts, this ability may transfer to situations where a user had a conceptual model for a website, they may be able to quickly scan through thumbnails, provided the thumbnails were of a sufficient resolution and size. Given this, these results have possible implications for visual bookmarks or the use of

thumbnails in browser interfaces as a way of displaying frequently visited websites, which are found in browsers such as Google Chrome, Apple Safari, etc.

Study Limitations

This study was limited by the stimulus resolutions and stimulus sizes. While there were several combinations of resolutions and sizes, it is undetermined whether the size and resolution interaction would occur with larger sizes or lower resolutions than what were used in this study. In terms of available visual information, there was a significant decrease moving from the Medium size to the Small size. There may be an intermediate size that leads to significant decreases in sensitivity. Additionally, lower resolutions should be tested in conjunction with the larger sizes of stimuli in the study to determine the minimum amount of information required for detection of websites by participants.

Conclusion

The results of this study demonstrated that participants relied on local features to detect websites during the RSVP tasks in the study for at least three of the four website categories. Stimulus size interacted strongly with the type of website, indicating that when features become indistinguishable where they cannot be processed meaningfully, sensitivity for detecting websites that depend on these features for identification decreases. This was the case for shopping websites, news websites, and social networks/blogs, but not for search websites. The interaction between stimulus size and resolution demonstrated that even when resolution decreases, as long as the image is of sufficient size to allow for meaningful processing of web elements, the decreases in resolution were not detrimental. However, if the stimulus size was small enough where there was a decrease in sensitivity at high resolutions,

the decrease in sensitivity was only increased as resolution decreased. Finally, changes in bias indicated that participants adjusted their criterion based on the amount of information available.

CHAPTER 3

OVERALL DISCUSSION AND CONCLUSION

Global-to-local processing has been a common theme within scene perception (Navon, 1977). Various scene perception theories have proposed that scene perception occurs first in terms of global features, which provides contextual information regarding the local features of the scene (Friedman, 1979; Henderson & Hollingsworth, 1999; Torralba, et al., 2006). The idea of global-to-local processing also has been supported in other perception and visual search theories, where global features and feed-forward mechanisms can be used to aid the processing of local features (Torralba, et al., 2006; Wolfe, et al., 2011). In essence, gist, or semantic information, is the result of the processing of these global features in such theories. In both RSVP studies, participants were capable of accurately detecting websites with above chance accuracy from stimulus exposures not exceeding the length of a fixation. In many cases, it was necessary for the participants to have additional processing time to do so. The above chance accuracy in both studies indicated that schemas or conceptualizations of websites provided information to detect websites from approximately 240 ms of stimulus exposure.

The SOAs required for website detection typically exceeded those for detecting upright and inverted natural scenes, which indicated that additional exposure time was needed, if not additional processing time. These differences were generally meaningful, as indicated by the moderate and large effect sizes. However, the lack of significant differences and small effect sizes between shopping and search websites with natural scenes indicated more efficient performance for processing websites. From discussions with participants, web elements may have aided detection. Participants stated they used specific web elements or diagnostic

features to aid in the detection of websites. The durations of the SOAs suggested that a few local features could likely be processed within this time. If these web elements were distinct or unique to the websites in question, detection of the website category could be based on detection of these elements or features instead of global features.

This was supported in the final study, where the size and resolution of websites were systematically manipulated. The results of the final study demonstrated that when the size of the stimulus decreased, sensitivity for detecting websites was diminished with the exception of search websites. Search websites typically lacked pictures and had other distinct attributes that appear to have been invariant to decreases in stimulus size. However, for the other three website categories (e.g., news, shopping, social networks/blogs), decreases in stimulus size had profound effects on sensitivity. Sensitivity dropped significantly for detection of news websites, shopping websites, and social networks/blogs when the smallest stimulus sized was displayed. The small stimulus sizes would have reduced the distinguishability of local features, such as pictures, text, or other web elements. This provided evidence that web elements were likely significant contributors to the detection of websites in this research.

In terms of resolution, the largest decreases in sensitivity were found when in the smallest stimulus size when resolution was decreased. Sensitivity for news websites, shopping websites, and social networks/blogs decreased to the extent where targets were indistinguishable from the distractor websites. This did not occur for search websites. Given these results, it appears that participants used local features as their primary factor for discriminating between the categories of websites for news websites, shopping websites, and

social networks/blogs, but search websites may have been detected due to either more distinct local features or their global features.

When significant differences were found for resolution and size, these differences were meaningful. Overall, large effect sizes were associated with significant differences in sensitivity between the Small and Medium/Large stimulus conditions for three of the four website categories. Similarly, large effect sizes were found with decreases in resolution for the Small stimulus condition when all the website categories were collapsed.

Commentary provided by participants in both studies demonstrated that they have specific schemas for the categories of websites used in this study. These schemas were typically based on the combinations of web elements found on websites. This included a variety of web elements such as pictures, headers, advertisements, search boxes, text, and links. Generally though, using a single element was a poor strategy for detection given the commonality of the elements. The most specific web elements were typically associated with search websites, and but not the other three categories. For instance, search websites were typically described as lacking pictures, containing a lot of white space, search boxes, advertisements, and having lists of blue hyperlinks. Combination of these attributes was not associated with the other categories of websites. Granted the other categories had hyperlinks, search boxes, contained white space, and even advertisements. However, the other website types typically contained pictures in combination of these other elements. Due to this, it appears that the lack of pictures in search websites or a combination of these attributes contributed to higher sensitivity in Low resolution and small stimulus sizes for search websites.

Such an idea also fits the data associated with the other three categories. The presence of pictures and the type of content appear to be integral to the website schemas of the four website categories in this study. Participants remarked on differences in picture content for shopping websites, news websites, and social networks/blogs. If local features, such as the pictures on these websites was diminished in quality, then that should affect the sensitivity of the websites as well. This was what occurred. As stimulus size increased, sensitivity also increased for the three types of websites that contained pictures. Website types that contained pictures were the most affected by decreases in both resolution and stimulus size. In Small size, High resolution stimulus conditions, participants were capable of distinguishing local features such as pictures. As the resolution decreased, quality of these local features also decreased. Moreover, as the stimulus size increased, the quality of the local features also increased. Additionally, the impact of decreases in resolution diminished as the stimulus size increased.

Shopping websites had more distinct pictures in comparison with news websites or social networks/blogs. Accordingly, as the size increased, sensitivity for shopping websites surpassed that of news and social networks/blogs. Additionally, the degree of bias also changed for shopping websites where decreases in resolution led to decreases in conservative bias.

Overall, this provides support that participants used local features or web elements, such as pictures and their content, for detection of news websites, shopping websites, and social networks/blogs in the study. The duration of the SOAs for each type of website would support the processing of a few local features, which was likely sufficient to detect targets during the RSVP streams. How search websites were detected was less clear. It was plausible

that both web elements found on websites, or the global features of search websites, could have contributed to their detection.

Theories of scene perception

Several theories of scene perception have been presented in the past, which typically reside into two different theoretical foundations. One of these foundations state that the context (semantic information) of a scene can be used in the recognition of objects, while the other states that object recognition can be completed separately or prior to the recognition of a scene (Friedman, 1979; Henderson & Hollingsworth, 1999; Torralba, et al., 2006). Models that use the context of a scene to aid in object recognition include the perceptual schema model, the priming model, and contextual guidance model, while the functional isolation theory posits the latter. Much support for contextual influences has been provided through literature (Davenport, 2007; Davenport & Potter, 2004; Potter, 1975; Potter, 1976; Rousselet, et al., 2005; Torralba, 2009). More recently, a visual search theory posited that these may occur in parallel with semantic information derived from scenes being used in parallel to aid in the processing of local features (Wolfe, et al., 2011).

The results of this research can be explained in terms of both positions. The results support functional isolation because the participants needed to rely on local features for website detection of at least three of the four website categories. For theories such as contextual guidance, the influence of global features during the detection tasks was not as readily apparent. The results suggest that schemas were activated, but the influence from spatial layout was a not primary factor for driving detection. In terms of theories driven by global information, differences between some genres of websites may be akin to the

differences between a front room and a bedroom. For website genres, the different genres may contain distinct semantic information similar to how both bedrooms and front rooms do, but the spatial structures of the genres are similar.

Research by Greene & Oliva (2009b) suggested that global features lacking distinct diagnostic qualities lead to decreased ability for discriminating targets from distractors. In this research, if website genres shared similar global features such as spatial layout, participants may have difficulty discriminating. When web elements were no longer distinguishable, this was what occurred for news, shopping, and social networks/blogs. This implies that the global features between these three types of websites were similar.

In terms of gist, the results of the studies provide some indication of the type of gist (perceptual or conceptual) used for detection. The time course suggested by Oliva (2005) states that perceptual gist leads to conceptual gist, and that perceptual gist contains global information such as spatial layout. For shopping, news, and social networks/blogs, the reliance on pictures and other local features was notable in the smallest size stimulus with lower resolutions. What appears absent from the information used for detection was spatial layout. This does not eliminate that gist was extracted from the websites, but it does suggest that the gist associated with various categories of websites (i.e., news, shopping, social networks/blogs) lacked sufficient, unique information regarding spatial layout, that would allow for detection. This suggests that participants were relying on conceptual gist to discriminate between the types of websites, which basically contains the semantic information regarding the genre. In accordance with the definition of perceptual gist proposed by Oliva, it appears that the

perceptual gist associated with genres would contain spatial layout information, but it was likely similar between the news, shopping, and social networks/blogs.

Website schemas

It is interesting that participants could detect category targets of websites with above chance accuracy with such brief exposures. This, in conjunction with the website classification study, strongly suggested that participants have schema representations of websites for these particular genres. Additionally, this suggests that such schemas are not necessarily unique to specific websites in the genre. For instance, this indicates that if a user has a conceptual model of a shopping website, their model would be applicable to multiple shopping websites instead of only a specific search website such as Amazon or eBay.

In terms of classification schemas, website categories fit well as basic categories in a superordinate-basic category relationship. Basic-level categories maximize within category similarity while maximizing out of category dissimilarities (Rosch, 1978). Essentially, websites classified as one classification share similar characteristics, while not sharing characteristics commonly found in other classifications of websites. Each website within this study was previously classified as exemplars for the respective type of category that maximized in-class agreement and minimized out-class agreement. For instance, several search websites shared common lists of results that consisted of hyperlinks, snippets, and URLs. Moreover, they also had search boxes in the center of the screen. Similarly, shopping pages were characterized by the presence of products, pages that allow selection of products, and also shopping cart elements such as the ability to change options regarding products, the presence of prices, etc. Social networks/blogs were characterized with profile pictures, snippets of text intermixed with

small pictures of people, and pages containing pictures of several people. News websites were typically found to have articles containing text and pictures, or dense combinations of text links and pictures intermixed.

In line with this taxonomy, it would place “websites” at a superordinate category, while the individual classifications of websites (search, shopping, news, and social networks/blogs) were considered basic categories. For each of these basic categories, naming specific examples was an easy task that required little effort for participants in latter studies. For instance, participants named Facebook or Twitter as examples of social networks/blogs; Google, Bing, and Yahoo! as examples of search websites; Amazon, EBay, and Wal-Mart as shopping websites; and CNN, BBC, and MSN as examples of news websites. Previous research has shown that superordinate categories and basic categories share similar features for scene perception and classification, which appears to be the case for the four types of websites in this study (Green & Oliva, 2009a; Oliva & Torralba, 2006).

If one examines web pages in terms of superordinate classification, web pages mostly consist of a combination of pictures, text, and other web elements that interconnect to other web pages, documents, or media. Due to this, it may be that webpages or websites themselves can be considered a basic category instead of the superordinate category and that the types of websites fall into subordinate categories. However, this taxonomy does not account for the ease of naming specific examples nor high degree agreement or the differences found between the genres of websites in this research.

First impressions in a glance

Judgments for visual appeal, trustworthiness, and perceived usability of websites can be determined from exposures that were 50 ms or shorter (Albert, et al., 2009; Lindgaard, et al., 2006; Lindgaard, et al., 2011; Tuch, et al., 2012; Thielsch, & Hirschfeld, 2012). In accordance with scene perception literature (Fei-Fei, et al., 2007; Greene & Oliva, 2009b), the presentation times of these studies suggested that global features contributed to these ratings. The current research found that web elements, such as pictures, aided in the detection of websites with the possible exception for search. Given this, it seems that judgments for visual appeal, trustworthiness, and perceived usability would be the result of processing different information than what was used for detecting websites in this research.

Impressions of aesthetics can be formed from a very short duration (i.e., 17 ms) and correlate with the consistency of how well websites matches a single genre (Tuch, et al., 2012). This research calls into question whether consistency of the websites for one type of genre was measured or possibly a measure of good design principles. If genre related information was extracted within 17 ms as implied, this suggests that processing of global features specific to the genre occurred. However, the current research demonstrated that users have conceptualizations for genres, but that in many cases, participants relied on web elements for distinguishing between the types of websites, indicating that such global features may not have been used in the discrimination between the websites.

The first impressions for visual appeal, trustworthiness, and perceived usability can be determined very quickly and independently of the type of website, but this does not eliminate whether an influence of the type of genre could affect such first impressions. Tuch, et al.,

(2012) demonstrated that the degree of conformance to a type of website and the visual complexity of the website was correlated with visual appeal. It is possible that this degree of conformance and complexity could vary within the different types of genres. For instance, the results suggest that discriminating search websites from the other website genres may not have relied on local features or that search websites were very consistent and different enough to allow for easier detection. This could indicate that conceptualizations for search websites were more homogenous in design when compared with other types of websites. For instance, search home pages should have low visual complexity given the presence of the single search bar surrounded by white space. Moreover, search results were commonly lists of results with little variability from website to website. These layouts were consistent between different search websites, which may have not been the case for the news websites, shopping websites, or social networks/blogs. As such, this may lead to differences in first impressions, such as visual appeal, between the genres of websites through latent factors such as prototypicality or visual complexity of the genre. Future research is needed to explore such possibilities between these first impressions and the various genres of websites.

Website conventions

The current research suggests that conceptualizations for some website types primarily contend with genre specific local features, and may not necessarily incorporate the spatial configuration of such features. This suggests that spatial constraints, such as those typically implicated in scene perception, may not be as rigid for websites as it pertains to how visual scenes are spatially represented.

Previous research showed that users quickly adapt to violations in website conventions. Many of these violations occurred by changing the placement of objects on a website (McCarthy, et al., 2003; Owens, et al., in press; Tzanidou, et al., 2005). By changing the placement, this may have led to the quick adaptations if spatial relationships between elements were not strongly represented in their conceptualizations. Having less specified spatial relationships between web elements may aid in the adaptation to violating website conventions. The quick adaptation to violations and the lack of rigidity in spatial relationships may be the result of the diverse array of websites encountered on the Internet. Even within each category, a wide variation of common elements, in various spatial arrangements, can be found. The varying degree of how these elements are arranged may impair our ability to develop strong representations for how websites are spatially arranged. However, this does not advocate for the violating best practices for design because such violations can have negative effects regarding usability, performance, and workload (Owens, et al., in press). However, web designers may have some latitude in the layout for news, shopping, and social networks/blogs that may not necessarily violate semantic conventions.

Research implications

The direct implications of this research are limited in the respect of what is perceived from websites within a few hundred milliseconds. How the findings of this research relate to how websites are perceived across multiple fixations or unlimited viewing needs further exploration. However, this research suggests that conceptual information regarding the website can be extracted from small screenshots and in some cases, with subpar resolution. These findings are directly applicable to interfaces that use thumbnails of websites, such as

smartphones or web browsers that display favorite or frequently visited websites. The SOAs, presentation times, stimulus sizes, and stimulus resolutions provide some guidance into the display quality and the display rate necessary to facilitate rapid scanning of thumbnails and how variations may affect user performance.

Previous exploration into the perception of websites lacked theoretical underpinnings based on current scene perception literature. Tapping decades of scene perception literature provides a foundation for exploring the conceptualizations of websites based on an established literature body. By having an established and solid theoretical foundation to build upon, the perception of websites can be explored systematically and in the realm of human perceptual processes. Such exploration provides theoretically-grounded contributions to human computer interaction and human factors, but also expands knowledge regarding scene perception into previously unexplored domains.

Based on scene perception literature, it appears that both global and local aspects of websites need to be taken under consideration during product development life cycles. This research, in conjunction with previous research for first impressions of websites, illustrate that a multitude of information, such as the quality of the website and other semantic information, can be perceived from websites within a glance. This information ranges from global characteristics of web pages (i.e., visual appeal) to specific elements found on web pages (i.e., pictures). Research has shown that the quick judgments are stable for extended viewing durations and it may be posited that the extracted semantic information can influence user behavior during extended interaction. Due to this, designers and developers need to consider perception and conceptualizations of websites from early vision to extended periods of

interaction. If designers and developers ignore early perceptual and conceptual aspects of websites, they may be handicapping the usability and interaction of the website in later interactions by ignoring information gleaned much earlier.

The extracted semantic information illustrates that website users have robust conceptualizations for various genres of websites. The conceptualizations for search, news, shopping, and social networks/blogs encompass various web page layouts that incorporate a variety of web page elements. These conceptualizations provide sufficient information allowing viewers to discriminate between the various types of websites based on a few hundred milliseconds of exposure. They are also flexible in allowing for differences in layout and the types of web page elements that may be encountered across the genres. Due to the presence of multiple websites found within the genres, these conceptualizations are not limited to only websites that participants have had previously viewed, either. This information can be useful from a web design perspective as it illustrates genres have generalizable characteristics, which can be used in the development of researched-based best practices for various genres of websites. Given that users have conceptual models for genres of websites, there may be performance and usability advantages for adhering to such models. Such ideas have found support through decreased usability associated with violating various website design conventions (Owens, et al., in press).

The differences found between search websites and other types of websites possibly have future implications for search. A common search results page commonly contains a list of results with a title, snippet, and URL. This format accounted for all the search results pages within this research. Users exhibited significant sensitivity for discriminating these search pages

(and search home pages) when viewing severely degraded screenshots in terms of both size and resolution, which possibly suggests the usage of global properties. Substantial changes to the global properties through new interfaces (i.e., the grid format used within Microsoft Windows 8, tablets, etc.) may not match the global properties provided by conceptual models accessed within these studies. Due to not matching users' conceptualizations for search, they may have difficulty adjusting to the novel layout or exhibit a preference for search engines that fit their known conceptual models.

Research limitations

There are limitations to this research. A possible limitation of this research was the repeated use of distractors. While both RSVP studies recycled distractors, the targets for each type of website were used only once. No feedback was provided to participants on whether they responded correctly or not in a given trial. Due to this, learning effects over the course of the RSVP tasks would not be based upon performance-based feedback or repeated exposures to the same targets. This does not predicate that participants did not change criterion for their decision making based on familiarity. While the exact cause of such change would be unknown, it would not be based on experimental feedback or repeated targets. Previous research noted that visually dissimilar, but conceptually similar decoy scenes resulted in increased false alarms when compared to other conceptually dissimilar distractors (Potter, Staub, & O'Conner, 2004). Due to this, in order to exclude learning effects based on familiarity, conceptually and visually distinct distractors would need to be used in the study. Given this, the limitation associated with the repeated use of distractors would be limited to the increased familiarity of the

distractor, and not necessarily the participants learning the type of distractor through any feedback.

Another possible limitation of the study is users' prior experience with the websites in the studies. Even though the websites were debranded, users may have been familiar with some websites based on previous experience. For instance, this would be expected for websites such as Facebook, Google, and Amazon given their prevalence in the commonly visited websites. However, the impact of this familiarity may have been partially controlled and balanced through debranding and by having multiple exemplars for each genre.

A final limitation of the research was the number of genres examined. In both RSVP studies, only four website genres were examined. The Internet contains a variety of different genres which were not examined in these studies. For instance, during the classification study, map driven websites tended to be classified distinctly into the maps category, but were not used in the RSVP studies. While differences between the genres used in these studies were associated with moderate and large effect sizes, it is unknown if similar trends would be found with other types of genres that were less well-known or less distinct. Given this, the findings of this research should not be generalized past the genres within this study without future research being conducted.

Future research

Previous studies have explored violations of website conventions by moving web elements to unexpected locations. While this changes the spatial structure of the web page, this does not examine semantic inconsistencies of the web elements. Given the results of this research, it indicates that the spatial relationship between web elements may be flexible and

partially responsible for the quick adaptation to changes in the layout of web elements. Future research should explore the implications of introducing semantically inconsistent local features within websites, particularly for pictures. Through the introduction of semantic inconsistencies that are not based on spatial structure, it may be possible to further explore user reliance on such features. For website types such as news, shopping, or social networks/blogs, a possible way to examine these effects would be to introduce pictures that are genre neutral (i.e., not related to any type of website) or that are genre inconsistent (i.e., replace product pictures with pictures from news).

The descriptions provided to participants in this study were basic categories. Future research should explore how richer descriptions or exemplar images of the website stimuli affect detection of stimuli. For example, examination of varying classification systems could be beneficial. This study focused on basic categorical classification derived from the primary purpose of the websites. In this classification system, individual page types were collapsed into semantically different basic categories. There may be a benefit for using richer descriptions of the website targets such as “search results” instead of “search,” or “product page” instead of “shopping,” or “news article” instead of “news.”

Research regarding first impressions of websites should be further explored in terms of genres of websites. Previous research exploring visual appeal may have been correlated with good design rather than being correlated with how well a website adheres to mental models of a website genre. By examining first impressions for the various genres, it may be possible to determine if differences exist in the time course for genre-related first impressions.

Finally, this research provides some information to designers about how websites are initially perceived and what type of information is expected and accessed within the first few hundred milliseconds of viewing. This information can be expanded through future research for longer viewing periods and to examine other semantic information than the type of website, which can be used to inform designers about how websites are conceptualized and how conceptualizations vary between genres. For instance, this could be expanded to individual page types within each type of genre. Additionally, genres, such as search, need to be examined with other novel layouts, such as those found within Microsoft Windows 8.

Conclusion

This research has shown that participants can detect websites accurately when they were displayed for less than a fixation and when the participants were allowed additional processing time. Subjective comments and SOAs suggested that participants likely relied on local features for the detection of website targets for several website categories. This notion was supported when the size and/or resolution of stimuli were decreased to the extent that web elements were indistinguishable. Participants were unable to distinguish targets from distractors for news, shopping, and social networks/blogs, but not for search. The SOAs associated with websites were longer than those for upright and inverted scenes, indicating that participants needed additional stimulus exposure and/or processing to achieve similar performance. At best, the lack of significant differences between natural scenes, shopping and search websites indicated that similar processing may have occurred. However, this was likely not the case for shopping websites given the results of the subsequent study where participants could not discern between targets and distractors when displayed as Small size,

Low resolution images. However, for search websites, detection could have relied on either global or local features.

This research shows that categorical based schemas were readily accessible and sufficient to detect websites based on exposures not exceeding the duration of a fixation. Additionally, categorical-based schemas are likely not a factor for making judgments for visual appeal, trustworthiness, or perceived usability of websites. This is primarily due to the reliance on web elements during detection tasks and the short exposures used during the first impression studies. These schemas do contain a wide variety of information, which accommodates multiple layouts, multiple websites, and varying web elements. Finally, the results of this research may provide some insight into why users can adapt quickly to violations in website conventions. In contrast with scenes, schemas for websites may not have similar spatial relationships between objects that contribute to the genres of websites in a similar manner to how spatial relationships between objects contribute to the classification of scenes.

REFERENCES

REFERENCES

Albert, W., Gribbons, W., & Almadas, J. (2009). Pre-conscious assessment of trust: a case study of financial and health care web sites. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 53(6), 449-453.

Feldman, J. (2011). Computers and Thought—The Back Story. *Annals of the History of Computing, IEEE*, *33*(4), 56-61.

Alexa (2013). *Top sites: the top 500 sites on the web*. Retrieved on January 6, 2013 from http://www.alexa.com/topsites/global

Amazon (2012). *Amazon Mechanical Turk Requester*. Retrieved from https://requester.mturk.com

Bernard, M. L. (2001). Developing schemas for the location of common web objects. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 45*(15), 1161-1165.

Bernard, M. L. (2003). Examining user expectations for the location of common e-commerce web objects. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 47*(11), 1356-1360.

Bernard, M., & Sheshadri, A. (2004). Preliminary examination of global expectations of users' mental models for e-commerce web layouts. *Usability News*, *6*(2).

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology*, *14*(2), 143-177.

Blog. (n.d.). In *Merriam-Webster Dictionary*. Retrieved from http://www.merriam-webster.com/dictionary/blog

Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, *13*(1), 210-230.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433-436.

Breakwell, G.M., Hammond, S., Fife-Schaw, C., & Smith, J. A. (2006). *Research Methods in Psychology, 3rd ed.* Los Angeles: Sage Publications.

Bulmer, M. G. (1979). *Principles of Statistics, 2nd ed.* New York: Dover Publications.

Chittaro, L. (2006). Visualizing information on mobile devices. *Computer*, *39*(3), 40-45.

comScore (2012). *Black Friday billion: Kick-off to brick-and-mortar shopping season surges past $1 billion in e-commerce spending for the first time*. Retrieved from http://www.comscore.com/Insights/Press_Releases/2012/11/Black_Friday_Billion_Kick-Off_to_Brick-and-Mortar_Shopping_Season_Surges_Past_1_Billion_in_E-Commerce_Spending_for_the_First_Time

Cousineau, D. (2005). Confidence intervals in within-subjects designs: A simpler solution to Loftus and Masson's method. *Tutorial in Quantitative Methods for Psychology, 1*(1), 42-45.

Crowston K., & Williams, M. (2000). Reproduced and emergent genres of communication on the World Wide Web. *The Information Society, 16*(3), 201-215.

DeCarlo, L. T. (1997). On the meaning and use of kurtosis. *Psychological Methods*, *2*(3), 292-307.

Di Nocera, F., Capponi, C., & Ferlazzo, F. (2004). Finding geometrical associations between meaningful objects in the web: a geostatistical approach. *PsychNology Journal*, *2*(1), 84-98.

Diamond, R., & Carey, S. (1986). Why faces are and are not special: An effect of expertise. *Journal of Experimental Psychology: General*, *115*(2), 107-117.

Dillon, A., & Gushrowski, B. A. (2000). Genres and the Web: is the personal home page the first uniquely digital genre?. *Journal of the American Society for Information Science*, *51*(2), 202-205.

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General*, *121*(3), 275-277.

Dvorine, I. (1963). *Pseudo-Isochromatic Plates.* Baltimore, MD: Waverly Press, Inc.

Dye, M. W. G., Green, C. S., & Bavelier, D. (2009). The development of attention skills in action video game players. *Neuropsychologia*, *47*(8-9), 1780-1789.

Edmonds, R., Guskin, E., Rosenstiel, T., & Mitchell, A. (2012). *The state of the news media 2012: an annual report on American journalism*. Retrieved from The Pew Research Center's Project for Excellence in Journalism website: http://stateofthemedia.org/2012/newspapers-building-digital-revenues-proves-painfully-slow/newspapers-by-the-numbers/

eMarketer (2012). *Apparel drives US retail ecommerce sales growth.* Retrieved from http://www.emarketer.com/newsroom/index.php/apparel-drives-retail-ecommerce-sales-growth/

Epstein, R. A., Higgins, J. S., Parker, W., Aguirre, G. K., & Cooperman, S. (2006). Cortical correlates of face and scene inversion: a comparison. *Neuropsychologia*, *44*(7), 1145-1158.

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: is it really attention free?. *Journal of Experimental Psychology-Human Perception and Performance*, *31*(6), 1476-1491.

Facebook (n.d.). Timeline. *Facebook*. Retrieved from http://newsroom.fb.com/Timeline

Fidell, F. S., & Tabachnick, B. G. (2004). Preparatory data analysis. In Schinka, J. A., Velicer, W. F., & Weiner, I. B. (Eds.), *Handbook of Psychology, Vol. 2* (pp. 115-141). John Wiley & Sons, Inc.

Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene?. *Journal of Vision*, *7*(1), 1-29.

Li, F. F., VanRullen, R., Koch, C., & Perona, P. (2002). Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences*, *99*(14), 9596-9601.

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science*, *18*(10), 850-855.

Field, A. (2009). *Discovering Statistics using SPSS, 3rd ed.* Los Angeles: Sage Publications.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*(3), 316-355.

Granka, L., Hembrooke, H., & Gay, G. (2006). Location location location: Viewing patterns on WWW pages. In *Proceedings of the 2006 Symposium on Eye Tracking Research & Applications* (pp. 43). ACM.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*, 534-537.

Green, C. S., & Bavelier, D. (2007). Action-video-game experience alters the spatial resolution of vision. *Psychological Science*, *18*(1), 88-94.

Greene, M. R., & Oliva, A. (2009a). The briefest of glances The time course of natural scene understanding. *Psychological Science*, *20*(4), 464-472.

Greene, M. R., & Oliva, A. (2009b). Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cognitive Psychology*, *58*(2), 137-176.

Gundotra, V. (2012, December 6). Google+: communities and photos [Weblog Message]. Retrieved from http://googleblog.blogspot.com/2012/12/google-communities-and-photos.html

Harding, G., & Bloj, M. (2010). Real and predicted influence of image manipulations on eye movements during scene recognition. *Journal of Vision*, *10*(2), 1-17.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, *50*(1), 243-271.

Intraub, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(1), 1-12.

Intraub, H. (1981). Rapid conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(3), 604-610.

Intraub, H. (1984). Conceptual masking: The effects of subsequent visual events on memory for pictures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(1), 115-125.

Johnson, A., & Proctor, R. W. (2004). Attention: Theory and Practice. SAGE Publications, Incorporated.

Joubert, O. R., Rousselet, G. A., Fabre-Thorpe, M., & Fize, D. (2009). Rapid visual categorization of natural scene contexts with equalized amplitude spectrum and increasing phase noise. *Journal of Vision*, *9*(1).

Kelley, T. A., Chun, M. M., & Chua, K. P. (2003). Effects of scene inversion on change detection of targets matched for visual salience. *Journal of Vision*, *3*(1), 1-5.

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's New in Psychtoolbox-3? *Perception [Online]*, *36*, ECVP Abstract Supplement.

Knoche, H., McCarthy, J. D., & Sasse, M. A. (2005). Can small be beautiful?: assessing image resolution requirements for mobile TV. In *Proceedings of the 13th Annual ACM International Conference on Multimedia* (pp. 829-838). ACM.

Lam, H., & Baudisch, P. (2005). Summary thumbnails: readable overviews for small screen web browsers. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems* (pp. 681-690). ACM.

Larson, A. M., & Loschky, L. C. (2009). The contributions of central versus peripheral vision to scene gist recognition. *Journal of Vision*, *9*(10), 1-16.

Lindgaard, G., Dudek, C., Sen, D., Sumegi, L., & Noonan, P. (2011). An exploration of relations between visual appeal, trustworthiness and perceived usability of homepages. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *18*(1), 1-30.

Lindgaard, G., Fernandes, G., Dudek, C., & Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression!. *Behaviour & Information Technology*, *25*(2), 115-126.

Loftus, G. R., Shimamura, A. P., & Johnson, C. A. (1985). How much is an icon worth? *Journal of Experimental Psychology: Human Perception and Performance, 11*(1), 1-13.

Loftus, G. R., Nelson, W. W., & Kallman, H. J. (1983). Differential acquisition rates for different types of information from pictures. *The Quarterly Journal of Experimental Psychology*, *35*(1), 187-198.

Lo Loschky, L. C., & Larson, A. M. (2010). The natural/man-made distinction is made before basic-level distinctions in scene gist processing. *Visual Cognition*, *18*(4), 513-536.

Loschky, L. C., Hansen, B. C., Sethi, A., & Pydimarri, T. N. (2010). The role of higher order image statistics in masking scene gist recognition. *Attention, Perception, & Psychophysics*, *72*(2), 427-444.

Liu, H., Agam, Y., Madsen, J. R., & Kreiman, G. (2009). Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. *Neuron*, *62*(2), 281-290.

MacMillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide. 2nd ed*. Lawrence Erlbaum Associates: New Jersey.

McBryan, O. A. (1994). GENVL and WWWW: Tools for taming the web. In *Proceedings of the First International World Wide Web Conference, 341*.

McCarthy, J. D., Sasse, M. A., & Riegelsberger, J. (2004). Could I have the menu please? An eye tracking study of design conventions. In *Proceedings of HCI 2003*, 401-414.

McCullagh, D., & Broache, A. (2007, March). Blogs turn 10—who's the father? *CNET News*. Retrieved from http://news.cnet.com/2100-1025_3-6168681.html

Maniar, N., Bennett, E., Hand, S., & Allan, G. (2008). The effect of mobile phone screen size on video based learning. *Journal of Software*, *3*(4), 51-61.

MapQuest (n.d.). In *Encyclopedia Britannica online*. Retrieved from http://www.britannica.com/EBchecked/topic/1437190/MapQuest

Meng, M., & Potter, M. C. (2008). Detecting and remembering pictures with and without visual noise. *Journal of Vision*, *8*(9), 1-10.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology, 4*(2), 61-64.

Murphy, K., & Spencer, A. (2009). Playing video games does not make for better visual attention skills. *Journal of Articles in Support of the Null Hypothesis, 6*(1), 1-20.

Nandakumar, C., & Malik, J. (2009). Understanding rapid category detection via multiply degraded images, *Journal of Vision, 9*(6), 1-8.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9*(3), 353-383.

Neisser, Ulrich. (1967). *Cognitive Psychology*. New York: Appleton, Century, Crofts.

Nielsen (2012). *State of the media: U.S. digital consumer report, Q3-Q4 2011*. Retrieved from http://www.nielsen.com/us/en/reports/2012/us-digital-consumer-report.html

Oliva, A. (2005). Gist of the scene. *Neurobiology of Attention, 696*, 251-256.

Oliva, A., & Schyns, P. G. (1997). Coarse blogs or fine edges? Evidence that information diagnosticity change the perception of complex visual stimuli. *Cognitive Psychology, 34*, 72-107.

Oliva, A., & Schyns, P. G. (2000). Diagnostic colors mediate scene recognition. *Cognitive Psychology, 41*(2), 176-210.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International Journal of Computer Vision, 42*(3), 145–175.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research, 155,* 23-36.

Oliva, A., & Torralba, A. (2007). The role of context in object recognition. *TRENDS in Cognitive Science, 11*(12). 520-527.

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.

Owens, J. W., Palmer, E. M., & Chaparro, B. S. (in press). The pervasiveness of text advertising blindness. *Journal of Usability Studies*.

Pelli, D. G. (1987) The ideal psychometric procedure, *Investigative Ophthalmology & Visual Science, 26*, 366.

Pelli, D. G. (1997) The VideoToolbox software for visual psychophysics: Transforming numbers into movies, *Spatial Vision, 10*, 437-442.

Phan, M. H., Jardina, J. R., Hoyle, S., & Chaparro, B. S. (2012). Examining the Role of Gender in Video Game Usage, Preference, and Behavior. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 56*(1), 1496-1500.

Potter, M. C. (1975). Meaning in visual search. *Science, 187*, 965-966.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2*(5), 509-522.

Potter, M. C., & Fox, L. F. (2009). Detecting and remembering simultaneous pictures in a rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 35*(1)*,* 28-38.

Potter, M. C., Staub, A., & O'Conner, D. H. (2004). Pictorial and conceptual representation of glimpsed pictures. *Journal of Experimental Psychology: Human Perception and Performance, 30*(3)*,* 478-489.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*(8), 1457-1506.

Rehm, G. (2002). Towards automatic Web genre identification: a corpus-based approach in the domain of academia by example of the Academic's Personal Homepage. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences* (pp. 1143-1152). IEEE.

Ryan, T., Field, R. H., & Olfman, L. (2002). Homepage genre dimensionality. In *Proceedings of the Eighth Americas Conference on Information Systems (pp. 1116-1128), Dallas, TX*.

Roth, S., Schmutz, P., Pauwels, S., Bargas-Avila, J., Opwis, K., (2010). Mental models for web objects: Where do users expect to find the most frequent objects in online shops, news portals, and company web pages?. *Interacting with Computers, 22*(2), 140-152.

Rosch, E. (1978). Principles of categorization. In *E., Rosch, & B., Lloyd (Eds.), Cognition and Categorization* (pp. 27-48). Hillsdale, NJ: Lawrence Erlbaum.

Rousselet, G. A., Joubert, O. R., & Fabre-Thorpe, M. (2005). How long to get to the "gist" of real-world natural scenes. *Visual Cognition, 12*(6), 852-877.

Santa-Maria, L., & Dyson, M. C. (2008). The effect of violating visual conventions of a website on user performance and disorientation: how bad can it be?. In *Proceedings of the 26th annual ACM International Conference on Design of Communication* (pp. 47-54). ACM.

Santini, M. (2006). Identifying genres of web pages. In *Proceedings of TALN 2006*.

Santini, M. (2007). Characterizing genres of web pages: Genre hybridism and individualization. In *Proceedings of the 40th Annual Hawaii International Conference on System Sciences* (pp. 71-71). IEEE.

Shaikh, A. D., Chaparro, B. S., & Joshi, A. (2006). Indian Users' Expectations for the Location of Web Objects on Informational Websites. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50*(17), 1922-1926.

Shaikh, A.D., & Lenz, K. (2006). Where's the search? Re-examining user expectations of web objects. *Usability News, 8.1*.

Shore, D. I., & Klien, R. M. (2000). The effects of scene inversion on change blindness. *The Journal of General Psychology, 127*(1), 27-43.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological monographs: General and applied, 74*(11), 1.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers, 31*(1)*,* 137-149.

Stevens, J. P. (2009). *Applied Multivariate Statistics for the Social Sciences, 5th ed.* New York: Taylor & Francis.

Tabachnick, B. G., and Fidell, L. S. (2007). *Using Multivariate Statistics*, *5th ed.* Boston: Allyn and Bacon.

Taylor, G., Singer, M. J., & Jerome, C. J. (2009). Development and evaluation of the game-based performance assessment battery (GamePAB) and game experience measure (GEM). In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *53*(27), 2014-2018.

Thielsch, M. T., & Hirschfeld, G. (2012). Spatial frequencies in aesthetic website evaluations – explaining how ultra-rapid evaluations are formed. *Ergonomics, 55*(7), 731-742.

Torralba, A. (2009). How many pixels make an image. *Visual Neuroscience, 26*(1)*,* 123-131.

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network: Computer Neural Systems, 14*(3), 391-412.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features on object search. *Psychological Review, 113*(4), 766-786.

Tuch, A. N., Presslaber, E. E., Stocklin, M., Opwis, K., Bargas-Aliva, J. A. (2012). The role of visual complexity and prototypicality regarding first impression of websites: Working towards understanding aesthetic judgments. *International Journal of Human-Computer Studies, 70*(11), 794-811.

Tversky, B., & Hemenway, K. (1983). Categories of environmental scenes. *Cognitive Psychology, 15*(1), 121-149.

Tzanidou, E., Petre, M., Minocha, S., & Grayson, A. (2005). Combining eye tracking and conventional techniques for indications of user-adaptability. In *Human-Computer Interaction-INTERACT 2005* (pp. 753-766). Springer Berlin Heidelberg.

Wang, T. (2012, December 18). There are now more than 200M monthly @twitter users [Tweet]. Retrieved from https://twitter.com/TonyW/status/281056150651293697

Warton, D. I., & Hui, F. K. (2011). The arcsine is asinine: the analysis of proportions in ecology. *Ecology, 92*(1), 3-10.

Watson, A. B., & Pelli, D. G. (1983) QUEST: a bayesian adaptive psychometric method. *Perception & Psychophysics, 33*(2)*,* 113-120.

Webley, Kayla (2010). *A brief history of online shopping.* Retrievedfrom http://www.time.com/time/business/article/0,8599,2004089,00.html

Weekend City Press Review (2013). *Who, what, why, and when.* Retrieved from http://www.weekendcitypressreview.co.uk/about

Wickens, T. D. (2001). *Elementary Signal Detection Theory.* Oxford University Press: New York.

Wolfe, J. M., Võ, M. L., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences, 15*(2)*,* 77-84.

Zimmermann, E., Schnier, F., & Lappe, M. (2010). The contribution of scene context on change detection performance. *Vision Research, 50*(20), 2062-2068.

Xiao, J., Hays, J., Ehinger, K. A., Oliva, A., & Torralba, A. (2010). Sun database: Large-scale scene recognition from abbey to zoo. In *2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* (pp. 3485-3492). IEEE.

APPENDICES

APPENDIX A

WEBSITE LISTING

TABLE 32

NEWS WEBSITES USED AS STIMULI IN THE CLASSIFICATION STUDY

Websites	
ABC News	NBC News
Aljazeera	New York Post
BBC News	New York Times
Boston Globe	Politico
CBS News	Reuters
CNBC News	San Francisco Chronicle
CNET News	San Francisco Weekly
CNN News	The Atlantic
Daily Mail	The Onion
Economist	The Guardian
Engadget	The Telegraph
Fox News	USA Today
Jerusalem Post	Wall Street Journal
Wichita Eagle	Washington Post
KAKE News	Washington Times
Los Angeles Times	Yahoo News
MSN News	

TABLE 33

SHOPPING WEBSITES USED AS STIMULI IN THE CLASSIFICATION STUDY

Websites	
American Eagle	Macy's
Amazon	Neiman Marcus
Apple	Newegg
Bass Pro	Nike
Best Buy	Nordstrom
Bloomingdales	Old Navy
Costco	Overstock
EBay	Photojojo
Etsy	REI
Gap	Sam Club
Garmin	Sears
Hollister	Souq
HSN	Target
Ideeli	Thinkgeek
JC Penny's	Walmart
Kohl's	Wine
Living Social	Zappos
Lowes	

TABLE 34

SEARCH WEBSITES USED AS STIMULI IN THE CLASSIFICATION STUDY

Websites	
Ask	Infospace
AOL Search	Ixquick
Bing	Lexxe
Blekko	Linkup
ChaCha	Lycos
DeeperWeb	Mywebsearch
Dogpile	Omgili
DuckDuckGo	Reddif
Ecosia Search	Search
Eluta	Spokeo
Evi	Volunia
Goby	Webcrawler
Good Search	Wolfram Alpha
Google	Yahoo
Hotbot	Yandex
Indeed	Yippy
Info	Yummly

TABLE 35

SOCIAL NETWORKS USED AS STIMULI IN THE CLASSIFICATION STUDY

Websites	
American Eagle	Macy's
Amazon	Neiman Marcus
Apple	Newegg
Bass Pro	Nike
Academia	LinkedIn
Badoo	Meetup
Bebo	Myspace
Blackplanet	Netlog
Delicious	Orkut
Deviant Art	Pinterest
Facebook	Plaxo
Flickr	Quora
Friendster	Research Gate
Foursquare	StumbleUpon
Fubar	Tribes
Google Plus	Tumblr
Hi5	Twitter
LastFM	Where Are you Now?

TABLE 36

MAPPING WEBSITES USED AS STIMULI IN THE CLASSIFICATION STUDY

Websites	
Accuweather	Nokia
Bing Maps	Orbisterrum
Bullitt Agency	Skyvector
GeoCommons	Social Explorer
Google Maps	StreetMap
Michelin	Yahoo Maps
Navteq	Yvanrodic
NearMap	Zillow

APPENDIX B

STUDY I DATA SCREENING

The data was screened for outliers. In this classification study, participants typically classified a single stimulus into one or two categories. An outlier was classified as a participant, that on average, classified stimuli into 2 standard deviations higher than the mean number of classifications across all participants. Given this, the criterion was set at 1.85 mean classifications per stimuli. Based on this criterion, nine participants, or 3.31% of the participants, were excluded from data analysis. These participants categorized each stimulus into 2.45 (*SD* = .51) categories across their classifications. Typically, these participants appeared to lack effort and classified stimuli randomly or appeared to lack mental models for the classification categories. For instance, participants that appeared to lack mental models of websites would commonly classify news oriented web pages into 3 or more categories including as corporate websites, general knowledge websites, mapping websites, search engines, blogs, and social networks in addition to news websites. Likewise, they would classify search oriented web pages as corporate websites, general knowledge websites, mapping websites, blogs, news, and social networks in additional to search engines.

The percentage of participants that categorized each web page as one of the nine possible classifications was calculated. Normality was examined using histograms, kurtosis, and skewness. Significance testing for skewness and kurtosis with large sample sizes is problematic as the standard error is sensitive to larger sample sizes (Tabachnick & Fidell, 2007). Bulmer (1979) provides that skewness outside of ±1 is indicative of being highly skewed and values between ±1/2 and ±1 as being moderately skewed. Generally, all classifications were

moderately to highly skewed. Where a value of 0 is considered normal for kurtosis, the distributions for the different classifications had varying kurtosis where news (-.86) and shopping (-1.39) classifications had negative kurtosis, search (1.49) was platykurtic, and all other classifications (> 3.7), including social networking (3.86), was leptokurtic. Examination of the histograms was in agreement that the classifications departed from normality. Examination of Q-Q plots indicated that the distributions also lacked linearity. Given the violations of normality and linearity, confirmatory factor analysis was not appropriate for this data set.

APPENDIX C

STUDY 2 DATA SCREENING

In total, six participants did not either complete the study or were removed from further analyses. Three participants were screened due suboptimal near visual acuity. Two participants were removed due to not completing the experiment. One participant was removed from analysis due to poor performance.

In the study, performance of the participants was compared. Upon viewing the data, screening data based on performance across each category and accuracy threshold combination was not feasible due to the variability across all web and upright scene categories/accuracy threshold combinations. For each participant, their overall accuracy was calculated, which collapses across all accuracy thresholds and stimuli categories. Standardized scores (z-scores) were calculated. Standardized scores greater than 1.96 were considered outliers and were subsequently removed. Based on this criteria, one participant exceeded the cutoff for overall performance (M = 865.11 ms, z = 2.41). The participant, a 55-year old male, was retired and was borderline red-green color-limited. This participant was removed from subsequent analyses.

APPENDIX D

STUDY 2 EXTENDED PARTICIPANT DEMOGRAPHICS

Participants stated they use the computer and Internet on a weekly basis. See Table 37 for frequency usage of computers and the Internet.

TABLE 37

INTERNET AND COMPUTER USAGE (NUMBER OF PARTICIPANTS)

Time	Computer	Internet
1-10 hours	7	4
11-20 hours	5	8
21-30 hours	6	5
31-40 hours	1	2
More than 40 hours	0	0

Participants were asked what types of activities they performed while on the Internet. The most common activities included were email, entertainment, education, social networking, and communicating with others. Less common activities were shopping, surfing the web, gathering product information, and gaming. See Table 38 for Internet activity frequencies.

TABLE 38

INTERNET ACTIVITIES

Internet Activity	Frequency
E-mail	18
Entertainment	17
Education	16
Social Networking	16
Communicating with others (excluding email)	15
Shopping	13
Surfing the Internet	12
Gathering information for personal needs	10
Playing games	9
Gathering product information	9
Keeping updated on current events	7
Work/Business	6
Gathering travel information	4

Participants were asked about their video gaming experience. Overall, 5 of 19 participants did not play video games previously, while 10 participants stated they currently did not play video games at the time of the study. Four participants noted the most they have previously played video games was over 40 hours per week, nine stated they played 1-10 hours per week, and one noted playing 11-20 hours per week. For current video game players, seven participants played video games 1-10 hours per week. Two participants noted they currently play video games from 11-20 hours per week. Participants played a variety of video games. The most commonly played genres included puzzle/card games (*n*=9), action games (*n*=7), driving games (*n*=6), role-playing games (*n*=6), strategy games (*n*=7), and fighting games (*n*=7).

APPENDIX E:

POST EXPERIMENT QUESTIONNAIRE

1. Participant ID:
2. What is your age?
3. What is your gender?
    - Male
    - Female

4. Are you currently a student?
    - Yes, I am a full-time student
    - Yes, I am a part-time student
    - No, I am not a student

5. What is your current major?
6. What is your primary occupation?
    - Consulting
    - Education (not a student)
    - Engineering
    - Finance
    - IT
    - Legal
    - Management
    - Manufacturing
    - Media
    - Medical
    - Public Relations
    - Purchasing
    - Retail
    - Sales & Marketing
    - Securities
    - Student
    - I do not have an occupation
    - Other (please specify)

7. How often do you use a computer during the week?
    - I do not use a computer
    - 1-10 hours
    - 11-20 hours
    - 21-30 hours
    - 31-40 hours
    - More than 40 hours

8. How often do you use the Internet during the week?
    - I do not use the Internet
    - 1-10 hours
    - 11-20 hours
    - 21-30 hours
    - 31-40 hours
    - More than 40 hours

9. Select the reasons why you use the Internet:
    - Education
    - Shopping
    - Gathering product information
    - Entertainment
    - Work/Business
    - Communication with others (excluding e-mail)
    - Gathering travel information
    - Keeping updated on current news events
    - E-mail
    - Gathering information for personal needs
    - Games
    - Browsing or surfing the web
    - Social networking
    - Other (please specify)

10. How often do you use the following types of websites?

Scale: 1- Never 2 - Rarely 3 - Monthly 4 - Weekly 5 - Daily

- News websites (e.g., CNN)
- Search websites (e.g., Google)
- Shopping websites (e.g., Amazon)
- Social Networks/Blogs (e.g., Facebook)

11. Which news websites do you frequently visit?
12. Which search websites do you frequently visit?
13. Which shopping websites do you frequently visit?
14. Which social networking websites do you frequently visit?
15. How many hours per week do you currently play video games?
    - I do not play video games
    - 1-10 hours
    - 11-20 hours
    - 21-30 hours
    - 31-40 hours
    - More than 40 hours

16. What is the maximum number of hours per week you've ever played video games?
    - I do not play video games
    - 1-10 hours
    - 11-20 hours
    - 21-30 hours
    - 31-40 hours
    - More than 40 hours

17. How often do you play:

Scale: 1- Never 2 - Rarely 3 - Monthly 4 - Weekly 5 - Daily

- Action (e.g., Halo, Call of Duty)
- Simulation (e.g., The Sims, Spore)
- Driving (e.g., Forza, Mario Kart)
- Puzzle/Card (e.g., Tetris, Poker, Bejeweled)
- Role Playing (e.g. Elder Scroll, Fable, World of Warcraft)
- Adventure (e.g., Resident Evil, Grand Theft Auto)
- Strategy (e.g., Civilization, Starcraft)
- Sports (e.g., Madden, FIFA)
- Fighting (e.g., Soul Caliber, Mortal Kombat)
- Music/Dance (e.g., Guitar Hero, Just Dance)
- Education/Edutainment (e.g., Democracy, Math Blaster)
- Fitness (e.g., Wii Fit, Your Shape: Fitness Evolved)
- Social Network (e.g., Farmville, Empires & Allies)
- Retro/Classic (e.g., Pacman, The Original Donkey Kong)

18. Please rate your skill level for the following genres:

Scale: 1 - Novice 2 3 4 5 - Expert

- Action (e.g., Halo, Call of Duty)
- Simulation (e.g., The Sims, Spore)
- Driving (e.g., Forza, Mario Kart)
- Puzzle/Card (e.g., Tetris, Poker, Bejeweled)
- Role Playing (e.g. Elder Scroll, Fable, World of Warcraft)
- Adventure (e.g., Resident Evil, Grand Theft Auto)
- Strategy (e.g., Civilization, Starcraft)
- Sports (e.g., Madden, FIFA)
- Fighting (e.g., Soul Caliber, Mortal Kombat)
- Music/Dance (e.g., Guitar Hero, Just Dance)
- Education/Edutainment (e.g., Democracy, Math Blaster)
- Fitness (e.g., Wii Fit, Your Shape: Fitness Evolved)
- Social Network (e.g., Farmville, Empires & Allies)
- Retro/Classic (e.g., Pacman, The Original Donkey Kong)

APPENDIX F

STUDY INSTRUCTIONS

Today, you're going to be discriminating between different types of websites. These types of websites include:

- News sites
- Search engines
- Shopping sites
- Social networks/blogs

News websites include:

- News home pages
- News pages that contain different topics of news linking to new stories
- News stories/articles

Search websites include:

- Primary search engine pages
- Search engine results

Shopping websites include:

- Shopping home pages
- Department pages for different types of products linking to product pages
- Individual product pages

Social networks/blogs include:

- Member profile pages
- Streams or feeds of member posts
- Pages that display links and pictures to other member profiles

On the screen, you're going to be presented with a website category. The website category is representative of the website itself and not individual website elements. After the website category is displayed, there will be a rapid series of websites presented. You will be asked to determine whether the website category was present or not within the series of websites. You'll have 5 seconds to respond after the websites have been presented, so please respond quickly and go with your first instinct.

Today, you're going to be discriminating between natural scenes. The types of natural scenes you will be seeing today include:

- Beaches: a pebbly or sandy shore, especially by the sea between high- and low-water marks
- Mountains: a large natural elevation of the earth's surface rising abruptly from the surrounding level
- Deserts: a waterless, desolate area of land with little or no vegetation, typically one covered with sand
- Forests: a large area covered chiefly with trees and undergrowth

On the screen, you're going to be presented with a scene category. The scene category represents the primary theme of the scene. After the scene category is displayed, there will be a rapid series of scenes presented. You will be asked to determine whether the scene category was present or not within the series of scenes. You'll have 5 seconds to respond after the scenes have been presented, so please respond quickly and go with your first instinct.

APPENDIX G

GENERAL APPROACH FOR ANALYSES

Analysis of the data for normality and parametric assumptions indicated that the dataset deviated from normal distributions. Previous research noted that skewness tends to affect statistical tests dependent upon sample means, while kurtosis tends to affect variance analyses (DeCarlo, 1997). Due to this, nonparametric analyses were conducted when data significantly departed from normal skewness or kurtosis, and when nonnormality of data was noted with tests for normality. Additionally, data was visually inspected through histograms or Q-Q plots. The Shapiro-Wilk test for normality was recommended for being more robust and powerful than Kolmogorov-Smirnov tests (Field, 2009). It was also noted that parametric tests such as ANOVAs and *t*-tests may be considered robust to moderate violations of normality (DeCarlo, 1997). When grouped data (e.g., conditions, levels, etc.) has at least 20 degrees of freedom for error, normality assumptions are typically met (Fidell & Tabachnick, 2004).

Where appropriate, transformations were applied and parametric statistics were conducted. Holm-Bonferroni procedures were applied to both post hoc tests and planned comparisons to address increased familywise Type I error. Holm-Bonferroni corrections are stepwise corrections where the statistical tests with the smallest *p*-values receive the most stringent comparisons. Transformed means and mean differences were provided when transformations were used in analyses for comparison purposes. In these analyses, the original means were also provided.

APPENDIX H

STUDY 2 EXTENDED RESULTS

For upright and inverted scenes, performance measures for mountains and beaches were collected. An analysis was conducted to determine whether the mountain and beach scene conditions could be collapsed across their respective scene types. Collapsing across scene types allowed for more straightforward analyses due to having a single standard that was representative of gist and another single standard that was representative of reduced/degraded gist.

A 2x3 repeated measures ANOVA was conducted on log transformed SOAs to determine differences between upright mountains and beaches across the three accuracy thresholds. The main effect for scene type was not significant, $F(1, 18) = 3.10$, $p = .10$. The main effect for accuracy threshold was significant, $F(2, 18) = 74.51$, $p < .01$, partial $\eta^2 = .81$. There was no significant interaction between scene type or accuracy threshold, $F(2, 36) = .62$, $p = .54$. The SOAs increased significantly across the three accuracy thresholds, but no significant differences were noted between upright beach and mountain conditions. This main effect had a large effect size. While participants required shorter SOAs to discriminate upright beaches than upright mountains at both the 75% and 90% accuracy thresholds, these differences were not significant.

Wilcoxon sign rank tests were conducted comparing inverted beach and mountain conditions. No significant differences were noted between the between beach and mountain conditions overall, or at the 60%, 75%, and 90% accuracy thresholds, $p > .05$.

Overall, the upright conditions were weakly correlated and the inverted conditions were moderately correlated. Due to no significant differences between the conditions for each accuracy threshold, inverted beach and mountain conditions were collapsed into an inverted scene condition for the remainder of the analyses. Similarly, upright beach and mountain conditions were collapsed into an upright scene condition. A moderate effect size was found for 75% accuracy threshold of the upright scene comparisons. All other effect sizes were small. Figure 20 reflects the mean SOAs for upright and inverted scenes across the different accuracy thresholds. Table 39 contains the mean SOAs for inverted scenes and back-transformed SOAs for upright scenes. Table 40 contains the results of the statistical testing and relevant correlations for the comparisons of upright and inverted scene categories.

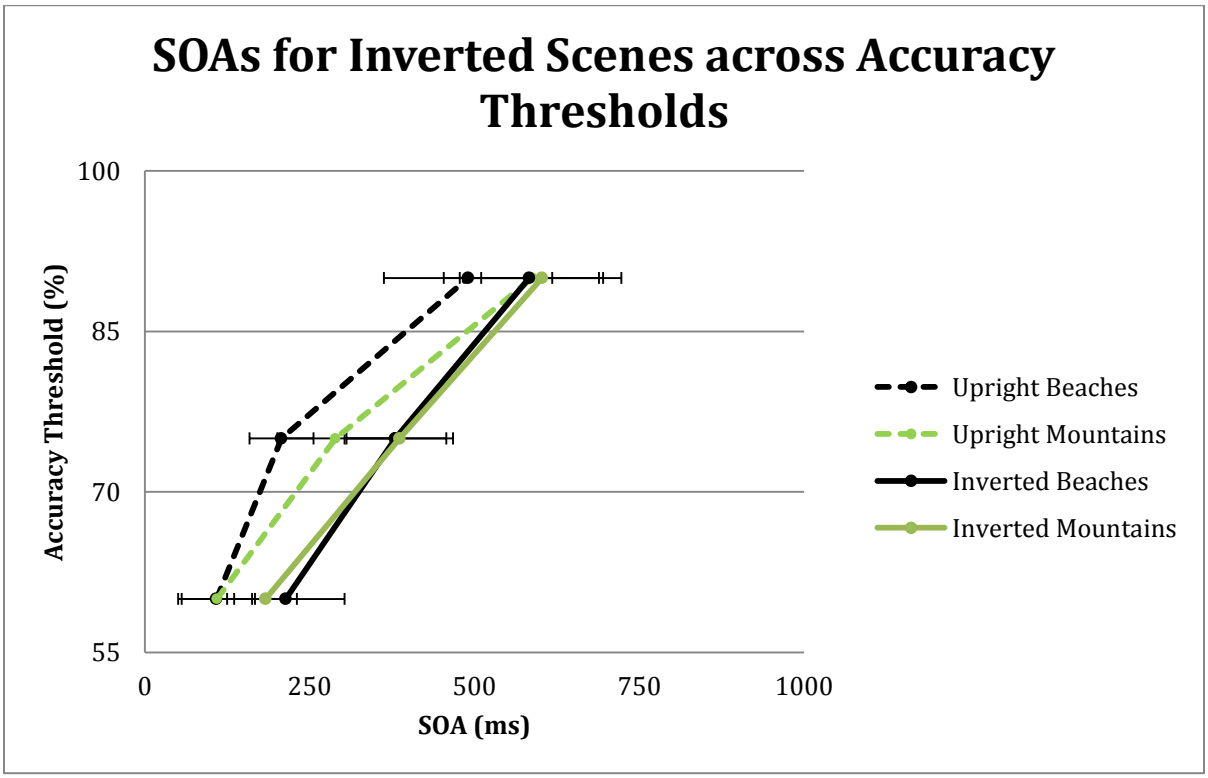

Figure 20. Mean SOAs for upright and inverted scenes across the accuracy thresholds.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

TABLE 39

MEAN SOAS FOR EACH NATURAL SCENE CATEGORY ACROSS ACCURACY THRESHOLDS

	Mean SOAs of inverted scenes and back-transformed SOAs of upright scenes (SD) or [CI] in milliseconds			
Category	60%	75%	90%	Overall
Beaches (U) (log)	100 [80, 125]	161 [134, 194]	384 [299, 494]	184 [159, 212]
Mountains (U) (log)	102 [83, 125]	217 [167, 282]	489 [348, 689]	221 [192, 256]
Beaches (I)	214 (165)	380 (263)	583 (245)	392 (157)
Mountains (I)	183 (173)	386 (279)	602 (290)	391 (219)

Note: Upright scenes are backtransformed values and within-subjects 95% Confidence Intervals (Cousineau, 2005; Morey, 2008). Inverted scenes are original values paired with standard deviations. (U) indicates the upright scene condition. (I) indicates the inverted scene condition.

TABLE 40

STATISTICAL TESTS COMPARING UPRIGHT AND INVERTED SCENES

Condition	Accuracy Threshold	*t*	*p*-value	r^2 (effect size)	correlation (*r*)	correlation *p*-value
Upright	60%	-.11	.91	< .01	-.20	.41
Upright	75%	-2.07	.05	.19	.62	< .01
Upright	90%	-.95	.36	.05	-.20	.42
Upright	Overall	-1.85	.08	.16	.21	.40
Condition	Accuracy Threshold	Z	*p*-value	r^2 (effect size)	correlation (*r*)	correlation *p*-value
Inverse	60%	-.68	.49	.01	.16	.52
Inverse	75%	-.28	.78	< .01	.63	.01
Inverse	90%	-.64	.52	.01	.25	.31
Inverse	Overall	-.16	.87	< .01	.61	.01

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

Website Usage

Participants were asked how often they frequent various types of websites. Only in the news websites category did participants indicate they never used that type of website. Across the four website categories, search websites were used most often, followed by social networks/blogs. For both of these categories, participants were very specific and had a limited number of search engines or social networks they use often. See Table 41 for usage means of how often the website genres were used.

TABLE 41

MEAN RATING OF HOW OFTEN WEBSITE CATEGORIES WERE USED

Category	Usage (frequency)	Frequently used websites (*n*<1)
News	Never=4 Rarely=7 Monthly=3 Weekly=2 Daily=3	CNN (6), Yahoo (4), MSN (2)
Search	Weekly=3 Daily=16	Google (17)
Shopping	Rarely=6 Monthly=5 Weekly=5 Daily=3	Amazon (5), EBay (4), Wal-Mart (2)
SN/Blogs	Rarely=1 Weekly=2 Daily=16	Facebook (16), Twitter (4)

The relationship between website category SOAs and the usage of the websites categories was examined. After adjusting the alpha level with a Holm-Bonferroni procedure, no correlations were significant. However, two website categories (search and social

networks/blogs) showed trends where increased website usage was related to shorter SOAs. However, this may be the result of a ceiling effect, where both categories had few participants that indicated less than daily usage. Additionally, while the correlations suggest increased usage resulted in decreased SOA durations, this was not indicative of increased performance as social networks/blogs generally had the longest SOAs of the four website categories. See Table 42 for correlations between website usage and SOAs.

TABLE 42

CORRELATIONS BETWEEN WEBSITE CATEGORY USAGE AND WEBSITE SOAS

Category	60%	75%	90%	Overall
News	r = .13, p =.61	r =.38, p = .11	r =.21, p =.40	r =.17, p =.48
Search	r = -.29, p =.23	r = -.29, p = .23	r = -.34, p =.15	r =-.34, p =.10
Shopping	r = .11, p =.65	r =.04, p = .88	r =.38, p =.11	r =.21, p =.38
SN/Blogs	r = -.11, p =.65	r = -.48, p = .04	r = -.35, p =.14	r =-.24, p =.32

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

The relationships between the number hours of Internet usage per week and the website category SOAs were analyzed. Similar to website usage, it would be expected that increased Internet usage may result in decreased SOAs across the different website categories. However, neither significant correlations nor trends were noted, $p > .05$.

Video Game Playing

The relationships between overall video game play per week and the website category SOAs were analyzed. No significant correlations were noted with amount of overall game play per week, $p > .05$. Out of the sample, 10 of the 19 participants indicated they were not current

video game players. The SOAs across both websites and scenes were examined across those that indicated currently they played video games and those who did not. The results indicated no differences in SOAs between those that currently play at least 1-10 hours or more a week and those that currently do not play video games, $p > .05$. An overall video game score was created by averaging responses for the mean of weekly video game play (current and overall), average amount of game play across genres, and the average level of skill per genre. Overall, video game play was not significantly correlated, $p > .05$.

The relationship between action video games play time was analyzed independently. Both action video game play time and skill level were not correlated with the results of the RSVP tasks, $p > .05$.

Previous research has indicated that video game play may enhance visual attention and visual cognition (Feng, et al., 2007; Green & Bavelier, 2003; Green & Bavelier, 2007; Dye, et al., 2009). In this study, it was found that performance during the detection tasks was not related to general video game playing behavior, action video game play frequency, or action video game skill.

Accuracy Rate

Due to the use of QUEST in the previous experiment to obtain 75% accuracy thresholds, accuracy rates obtained for similar size and stimulus quality of the four categories were compared. These comparisons determined whether the SOAs for the four categories were equivalent performance-wise between the four category types. The accuracy rate for the Medium size, High resolution conditions for each website category was calculated. The

accuracy rates were transformed with logit transformations. Normality assessment across the four transformed accuracy rates was conducted. According to the Shapiro-Wilks tests for normality, news websites and social networks/blogs deviated from normal distributions. However, social networks/blogs was significantly leptokurtic, $p < .01$. Fidell and Tabachnick (2004) noted that with Shapiro-Wilks tests are very sensitive and may indicate significance even when normality is present. Moreover, they stated that normality assumptions for univariate and multivariate analyses can typically be achieved when there are at least 20 degrees of freedom for error when the data is grouped. Due to these assumptions, parametric statistics were conducted. It was also noted that parametric tests such as ANOVAs and *t*-tests may be considered robust to moderate violations of normality and that test of means were more sensitive to departures in skewness, while analysis of variance was more sensitive to departures in kurtosis (DeCarlo, 1997).

A one-way repeated measures ANOVA was conducted on the logit transformed accuracy rates for the four website types in the Medium size, High resolution condition to determine if performance was similar across the website categories. The results indicated no significant differences in the detection accuracy rates across the four types of websites in the medium size, High resolution condition, $F(3,60) = 2.42$, $p = .08$. These results indicated that the SOAs for the four types of websites used during the RSVP trials were sufficient to achieve similar overall accuracy rates to the previous study. Table 43 shows the backtransformed logit means and confidence intervals in addition to the original means prior to transformations. Figure 21 shows the untransformed means for the four website categories.

TABLE 43

MEAN ACCURACY RATES FOR MEDIUM SIZE, HIGH RESOLUTION CONDITIONS

Website Category	Backtransformed Means	Untransformed Means
News Websites	76 [69, 82]	73 [14]
Search Websites	78 [69, 85]	73 [17]
Shopping Websites	79 [71, 85]	75 [15]
Social Networks/Blogs	68 [62, 74]	67 [13]

Note: 95% Confidence Intervals are provided for backtransformed means and standard deviations are provided for untransformed means.

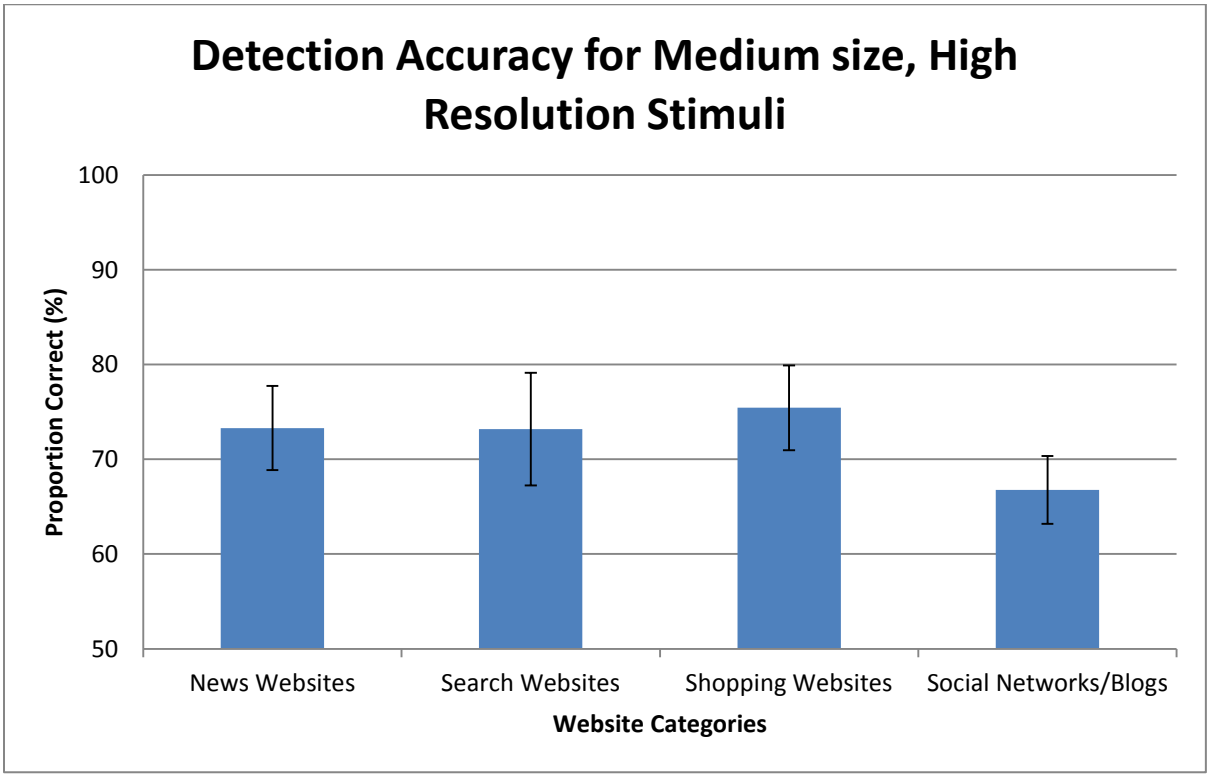

Figure 21. Accuracy for detecting websites in the Medium size, High quality condition.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

Additionally, for each website category, the overall accuracy rate was compared against the logit of 75% using one-sample t-tests. For news websites, the participants accuracy for detecting targets did not differ from the logit for 75%, $t(20)=.26$, $p = .80$. For search websites, the participants accuracy for detecting targets did not differ from the logit for 75%, $t(20)=.65$, $p = .53$. For shopping websites, the participants accuracy for detecting targets did not differ from the logit for 75%, $t(20)=1.00$, $p = .33$. For social networks/blogs, the participants accuracy for detecting targets did differ from the logit for 75%, $t(17)=-2.29$, $p = .03$. Overall, detection of Medium size, High resolution targets did not differ from each other, but accuracy detecting social networks/blogs was significantly lower than 75% accuracy.

APPENDIX I:

STUDY 3 DATA SCREENING

In total, eight participants were removed from the study. Out of the 29 participants, three participants were screened due to suboptimal near visual acuity, two were screened due to suboptimal color vision, two were excluded due to significantly higher nonresponse rates, and one participant was excluded due to not following instructions provided for the experiment. Overall, the remaining 21 participants were included in the analyses.

Participants were given five seconds to respond whether they thought the target was present or not. When participants did not respond within this timeframe, the response was marked as absent and incorrect. The overall response rate of the participants was analyzed. Participants had low nonresponse rates with a mean nonresponse rate of .35%. Individual responses were examined for outliers. Two participants had nonresponse rates that exceeded 2% with accompanying z-scores above 2.32. Due to the higher non response rate for both of these individuals, they were removed from further analysis. The removal of both of these individuals lowered the overall nonresponse rate to .14% or 17 trials. Of the participants kept in the analysis, the highest number of nonresponse trials was 7 out of 576 experimental trials.

APPENDIX J

STUDY 3 EXTENDED PARTICIPANT DEMOGRAPHICS

Of the 21 participants in the study, all participants had near visual acuity equal to 20/25 or better at 40cm and had normal color vision. Of these 21 participants, 20 were identified as being full-time students and 1 as a non-student. Thirteen participants stated that being a student was their primary occupation. The remaining participants worked as a bartender, as a law clerk, in public safety, in finance, in public relations, in education, in management, in embalming and crematory services, or were unemployed. Participant majors included psychology, criminal justice, engineering, nursing, health sciences, physics, exercise science, secondary education, and biology.

Participants stated they use the computer and Internet on a weekly basis. See Table 44 for frequency usage of computers and the Internet.

TABLE 44

INTERNET AND COMPUTER USAGE (NUMBER OF PARTICIPANTS)

	Number of Participants	
Number of hours	Computer	Internet
1-10 hours	8	7
11-20 hours	5	6
21-30 hours	2	4
31-40 hours	4	2
More than 40 hours	2	2

Participants were asked what types of activities they performed while on the Internet. The most common activities included were email, entertainment, education, social networking,

and communicating with others. Less common activities were shopping, surfing the web, gathering product information, and gaming. See Table 45 for Internet activity frequencies.

TABLE 45

INTERNET ACTIVITIES

Internet Activity	Frequency
E-mail	20
Education	19
Entertainment	16
Social Networking	15
Shopping	14
Surfing the Internet	12
Keeping updated on current events	11
Work/Business	9
Communicating with others (excluding email)	8
Gathering information for personal needs	8
Gathering product information	7
Playing games	4
Gathering travel information	5

Participants were asked about their video gaming experience. Overall, 5 participants have not played video games previously, while 9 participants stated they currently did not play video games at the time of the study. Two participants noted the most they have previously played video games was over 40 hours per week, eight stated they played 1-10 hours per week, two noted playing 11-20 hours per week and three participants stated they played 21-30 hours per week. For current video game players, 11 participants played video games 1-10 hours per week. Only one participant noted they currently play video games from 11-20 hours per week.

Participants played a variety of video games. The most commonly played genres included action games (*n*=8), puzzle/card games (*n*=7), driving games (*n*=6), and music/dance games (*n*=6).

APPENDIX K

STUDY 3 EXTENDED RESULTS

Web Usage

After the RSVP tasks, participants were asked how often they used news websites, search websites, shopping websites, and social networks/blogs. Overall, participants visited search sites and social networks/blogs most frequently. See Table 46 for the usage of each type of website and frequently visited websites.

TABLE 46

MEAN RATING OF HOW OFTEN WEBSITE CATEGORIES WERE USED

Category	Usage (frequency)	Frequently used websites (*n*<1)
News	Never = 4 Rarely = 5 Monthly = 3 Weekly = 7 Daily = 2	CNN (6), Kake News (4), Kansas.Com (3), BBC (2), Yahoo News! (2)
Search	Monthly = 1 Weekly = 5 Daily = 15	Google (19), Bing (6), Yahoo! (2)
Shopping	Never = 2 Rarely = 5 Monthly = 6 Weekly = 6 Daily = 2	Amazon (12), EBay (2), Barnes & Noble (2)
SN/Blogs	Never = 2 Rarely = 2 Monthly = 3 Weekly = 1 Daily = 13	Facebook (16), Twitter (4)

Note: * indicates a significant comparison.

The participants' sensitivity was compared with their website usage. Due to usage being an ordinal variable, Spearman correlations were conducted by stimulus type with the size and resolution conditions. Holm-Bonferroni adjustments were conducted to adjust for increased Type I familywise error. Across the four stimulus types, a negative relationship between social networking/blog usage and sensitivity was found for the Small size, Low resolution condition, $r = -.62$, $p < .01$, $r^2 = .38$. The negative relationship suggests that increased usage was related to decreases in sensitivity in the Small size, Low resolution condition. No other significant relationship was found between sensitivity and web usage. See Table 47 for correlations.

TABLE 47

CORRELATION RESULTS BETWEEN SENSITIVITY AND WEBSITE USAGE

Size	Resolution	News Websites		Search Websites		Shopping Websites		Social Networks/Blogs	
		r	*p*	*r*	*p*	*r*	*p*	*r*	*p*
Small	High	.25	.28	.29	.21	.14	.56	.22	.32
Small	Moderate	.28	.23	.26	.31	-.04	.85	-.24	.30
Small	Low	-.05	.83	.16	.48	-.19	.41	-.62	< .01*
Medium	High	-.35	.12	.34	.14	.28	.22	.03	.91
Medium	Moderate	.40	.07	.21	.36	-.03	.90	.18	.44
Medium	Low	.13	.56	.16	.47	.20	.38	.31	.17
Large	High	.16	.48	.35	.12	.21	.37	-.10	.66
Large	Moderate	.07	.76	.32	.16	.34	.14	-.23	.33
Large	Low	-.03	.91	.10	.68	.05	.83	.03	.89

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

Weekly computer and Internet usage was compared with the sensitivity. Spearman correlations were conducted. The results demonstrated no significant relationships for sensitivity with Internet usage.

Similarly to sensitivity, correlations with bias were conducted to determine if bias was related to previous web usage. Overall, no significant correlations were noted for any stimulus type or size and resolution combination. See Table 48 for correlations.

TABLE 48

CORRELATION RESULTS BETWEEN BIAS AND WEBSITE USAGE

Size	Resolution	News Websites		Search Websites		Shopping Websites		Social Networks/Blogs	
		r	*p*	*r*	*p*	*r*	*p*	*r*	*p*
Small	High	-.29	.20	.38	.10	-.06	.80	.21	.35
Small	Moderate	-.30	.19	.44	.04	-.17	.45	-.09	.70
Small	Low	-.14	.54	-.05	.82	.09	.69	-.07	.78
Medium	High	-.49	.03	.34	.13	-.28	.22	.03	.90
Medium	Moderate	-.19	.42	.30	.18	-.13	.57	.22	.33
Medium	Low	-.08	.74	-.05	.83	-.06	.80	-.01	.97
Large	High	-.54	.01	.07	.77	-.20	.38	.14	.54
Large	Moderate	-.37	.10	.07	.76	-.36	.11	.42	.06
Large	Low	-.48	.03	.27	.24	-.32	.17	.41	.07

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

Weekly Internet usage was compared with the bias. Spearman correlations were conducted. The results demonstrated no significant relationships for Internet usage and bias.

Video Game Playing

The relationships between overall video game play per week and participants' sensitivity was analyzed. Correlations were conducted between current game play and sensitivity. No significant correlations were noted with amount of current or overall game play per week with sensitivity. Composite gaming scores were create by averaging the mean genre game play frequency, mean skill level, and current and all time game play frequency. No significant correlations were noted with sensitivity. Out of the sample, 9 of the 21 participants indicated they were not current video game players. The sensitivity for each condition was examined across those that indicated currently they played video games and those who did not. The results indicated no differences in sensitivity between those that currently play at least 1-10 hours or more a week and those that currently do not play video games. Moreover, the relationship between action video game playing frequency and sensitivity was examined with correlations. No significant relationships were discovered. Finally, the skill level for those that play action video games was compared to the participants' sensitivity for each condition. Similarly, action game skill was found not to have a significant relationship with sensitivity for detecting websites.

Previous research has indicated that video game play may enhance visual attention and visual cognition (Feng, et al., 2007; Green & Bavelier, 2003; Green & Bavelier, 2007; Dye, et al., 2009). In this study, it was found that performance during the detection tasks was not related to video game playing behavior, action video game play frequency or action video game skill.

Bias Measure B"

To calculate B", the equation 3 was used.

$$B'' = sign(H - F)\frac{H(1-H) - F(1-F)}{H(1-H) + F(1-F)} \quad (3)$$

For the bias measure B", normality of the data was checked across each combination of website type (news, search, shopping, social networks/blogs), stimulus size (Small, Medium, Large), and stimulus resolution (High, Moderate, Low). Similar to B''_D, significant skewness and kurtosis was noted for select conditions. Significant kurtosis was noted for distributions of the Small sized, Moderate resolution stimulus condition of news, the Medium sized, High resolution condition for shopping stimuli, and the Moderate and High resolution versions of the Large social network/blog stimuli conditions. Distributions for the Medium size, High resolution shopping condition, and the Moderate and High resolution versions of the Large sized social networks/blogs stimuli.

A 3-way repeated measures ANOVA was conducted to examine the differences between the four website categories at the 3 stimulus sizes and 3 resolutions for B". The results of the repeated-measures ANOVA indicated significant main effects for stimulus size, $F(2,40) = 12.52$, $p < .01$, partial $\eta^2 = .36$. This main effect had a large effect size. The main effects for stimulus type $F(3,60) = 2.09$, $p = .11$, partial $\eta^2 = .10$, and stimulus resolution, $F(2,40) = 1.54$, $p = .23$, partial $\eta^2 = .07$, were not significant. The results indicated significant interactions also for stimulus type and resolution, $F(6,120) = 4.45$, $p < .01$, partial $\eta^2 = .18$, and stimulus size and resolution, $F(4,80) = 4.50$, $p < .01$, partial $\eta^2 = .18$. The effect sizes for both significant

interactions were large. The two-way interaction for stimulus type and size, $F(6,120) = 1.35$, $p = .24$, partial $\eta^2 = .06$, and the three-way interaction for stimulus type, stimulus size, and stimulus resolution were not significant, $F(12,240) = .58$, $p = .85$, partial $\eta^2 = .03$. Planned pairwise tests were conducted on the two two-way significant interactions. See Figure 22 for the interaction of stimulus resolution and stimulus type for B" and Figure 23 for the interaction between stimulus resolution and stimulus size for B".

The results of the planned comparisons for the interaction between stimulus type and resolution indicated participant's criterion was more conservative in the High resolution condition than either lower resolution condition for the shopping website stimulus condition. See Table 49 and 50 for the statistical tests of the planned comparisons.

The results of the planned pairwise comparisons for the interaction between stimulus type and stimulus size indicated that in the Moderate resolution, participants were more conservative when displayed Medium or Large sized stimuli than the Small stimuli. Moreover, in the lowest resolution, participants were more conservative when displayed Medium or Large sized stimuli than the Small stimuli. Finally, participants exhibited more conservative behavior when displayed Small stimuli in the highest resolution when compared to the lowest resolution. The effect sizes for the significant interactions were large. See Tables 51 and 52 for the statistical tests of the planned comparisons.

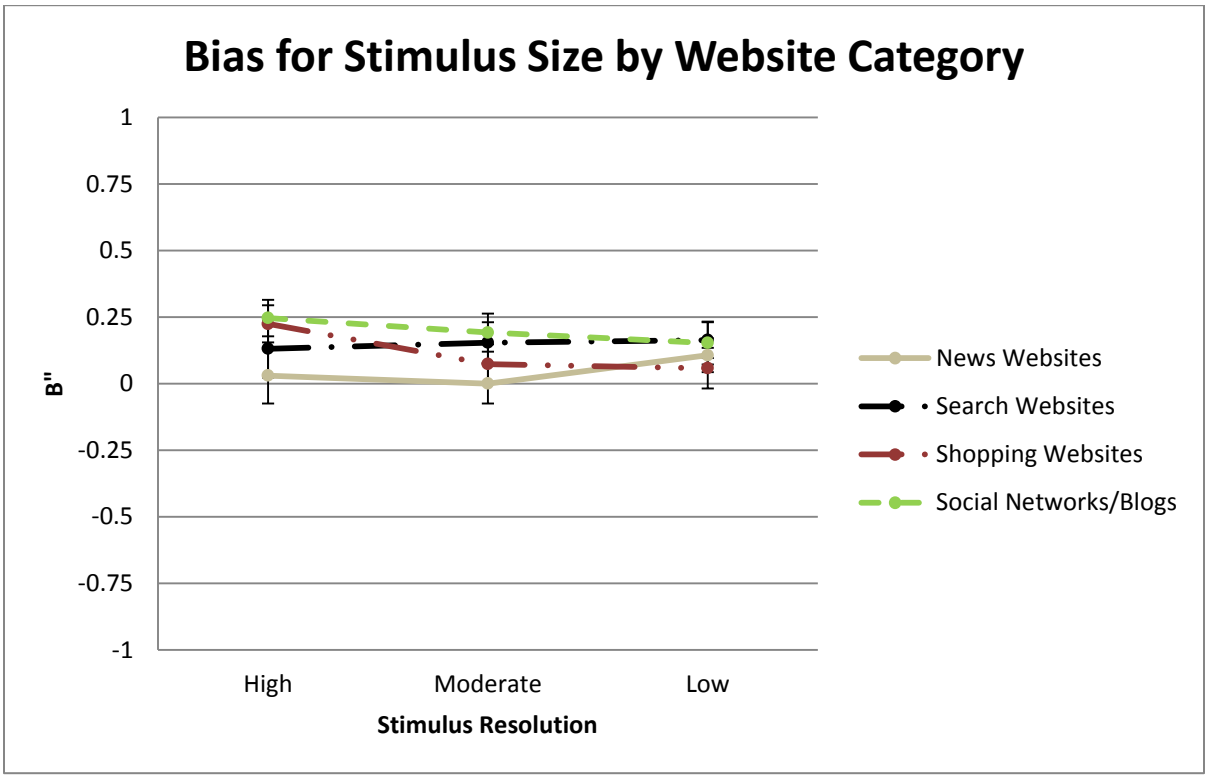

Figure 22. Bias for detecting websites across the resolution conditions.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

TABLE 49

STATISTICAL TESTS COMPARING STIMULUS TYPES AND STIMULUS RESOLUTIONS FOR BIAS B"

Comparison	*t*	*p*-value	r^2	MD
High: News Websites - Shopping Websites	-2.94	.01	.30	-.19
High: News Websites - Social Networks/Blogs	-2.80	.01	.28	-.22
Moderate: Shopping Websites - Social Networks/Blogs	-2.30	.03	.21	-.12
Low: Search Websites - Shopping Websites	2.02	.06	.17	.11
High: Search Websites - Social Networks/Blogs	-2.00	.06	.17	-.12
Moderate: News Websites - Social Networks/Blogs	-1.69	.11	.13	-.10
High: Search Websites - Shopping Websites	-1.59	.13	.11	-.10
Low: Shopping Websites - Social Networks/Blogs	-1.48	.15	.10	-.10
Moderate: Search Websites - Shopping Websites	1.40	.18	.09	.08
Moderate: News Websites - Search Websites	-1.16	.26	.06	-.06
High: News Websites - Search Websites	-1.09	.29	.06	-.10
Low: News Websites - Social Networks/Blogs	-1.02	.32	.05	-.05
Low: News Websites - Search Websites	-1.02	.32	.05	-.06
Low: News Websites - Shopping Websites	1.00	.33	.05	.05
Moderate: Search Websites - Social Networks/Blogs	-.62	.54	.02	-.04
High: Shopping Websites - Social Networks/Blogs	-.45	.66	.01	-.02
Moderate: News Websites - Shopping Websites	.39	.71	.01	.02
Low: Search Websites - Social Networks/Blogs	.23	.82	< .01	.01

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

TABLE 50

STATISTICAL TESTS COMPARING STIMULUS RESOLUTIONS BY STIMULUS TYPES FOR BIAS B"

Comparison	*t*	*p*-value	r^2	MD
Shopping Websites: High - Moderate	4.35	< .01*	.49	.15
Shopping Websites: High - Low	3.76	< .01*	.41	.17
Social Networks/Blogs: High - Low	2.19	.04	.19	.09
Social Networks/Blogs: High - Moderate	1.99	.06	.16	.06
News Websites: High - Moderate	-1.65	.12	.12	-.07
News Websites: High - Low	-1.90	.13	.11	-.08
Social Networks/Blogs: Moderate - Low	.89	.39	.04	.04
Search Websites: High - Low	-.69	.50	.02	-.03
Search Websites: High - Moderate	-.48	.64	.01	-.02
News Websites: Moderate - Low	-.34	.74	.01	-.01
Shopping Websites: Moderate - Low	.33	.74	.01	.02
Search Websites: Moderate - Low	-.22	.83	< .01	-.01

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

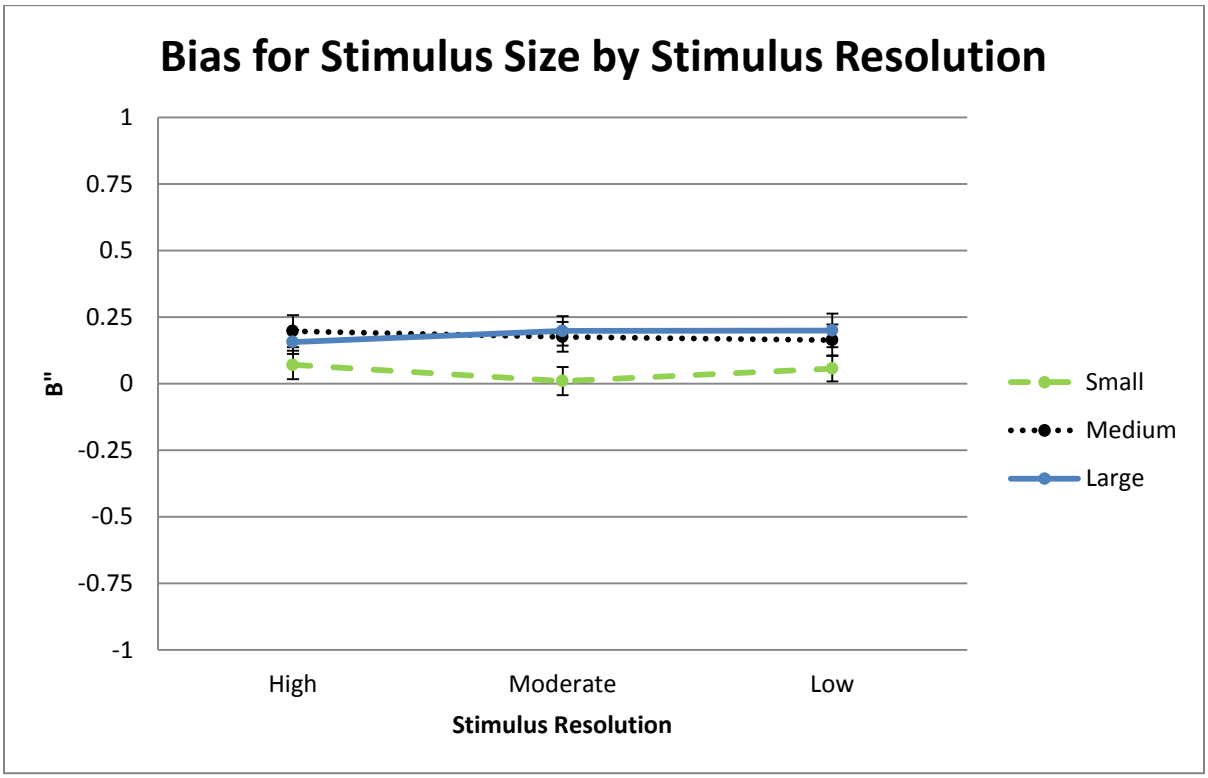

Figure 23. Bias across the stimulus resolution and sizes.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

TABLE 51

STATISTICAL TESTS COMPARING STIMULUS RESOLUTIONS BY STIMULUS SIZE FOR BIAS B”

Comparison	*t*	*p*-value	r^2	MD
Small: High - Low	3.55	< .01*	.39	.11
Small: High - Moderate	2.97	.01	.31	.10
Large: High - Moderate	-2.10	.05	.18	-.05
Medium: High - Moderate	1.10	.28	.06	.04
Medium: High - Low	1.08	.29	.06	.06
Large: Moderate - Low	.89	.38	.04	.03
Large: High - Low	-.68	.50	.02	-.02
Medium: Moderate - Low	.36	.73	.01	.02
Small: Moderate - Low	.05	.96	< .01	< .01

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

TABLE 52

STATISTICAL TESTS COMPARING STIMULUS SIZE BY STIMULUS RESOLUTION FOR BIAS B”

Comparison	*t*	*p*-value	r^2	MD
Moderate: Small - Large	-5.28	< .01*	.58	-.20
Low: Small - Large	-4.90	< .01*	.54	-.18
Low: Small - Medium	-4.22	< .01*	.47	-.14
Moderate: Small - Medium	-3.32	< .01*	.36	-.15
High: Small - Medium	-1.55	.14	.11	-.09
High: Small - Large	-1.32	.20	.08	-.05
Low: Medium - Large	-1.14	.27	.06	-.04
Moderate: Medium - Large	-1.09	.29	.06	-.05
High: Medium - Large	.64	.53	.02	.03

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

To assess how bias varied across the various conditions, B" was compared against 0, the value of B" that states no bias exists. One-sample *t*-tests were conducted for the conditions in the study. Bias for Social networks/blogs in the Medium and Large sized stimuli were conservative. Additionally, in the Small size, High condition of social networks/blogs, conservative bias was exhibited. Finally, conservative bias exhibited for shopping websites in the Large size, High and Medium size, High conditions. For the significant comparisons, large effect sizes were found. See Figure 24 for the results and Table 53 for the statistical test results.

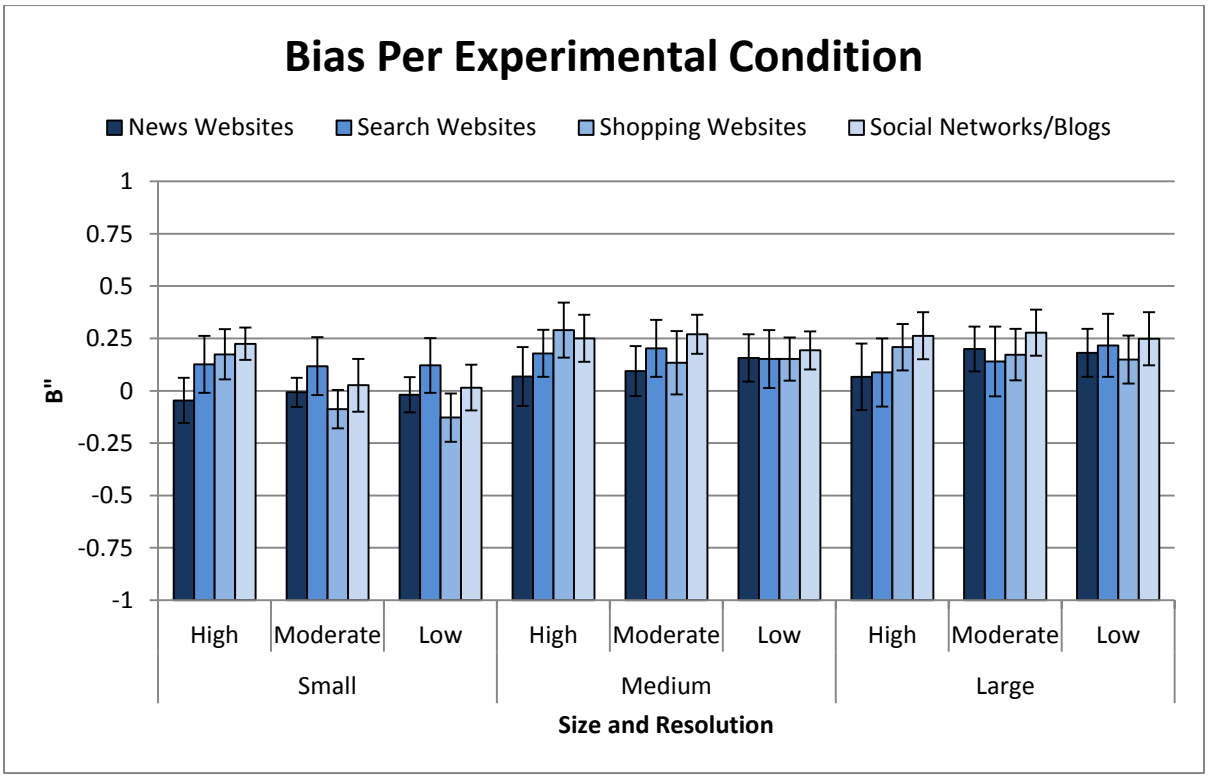

Figure 24. Bias for detecting websites in each of the conditions of the study.
Note: Error bars are within-subjects 95% confidence intervals (Cousineau, 2005; Morey, 2008)

TABLE 53

STATISTICAL TESTS COMPARING ALL CONDITIONS FOR BIAS B" AGAINST NEUTRAL BIAS.

Comparison	*t*	*p*-value	r^2	MD
Social Networks/Blogs: Medium - Moderate	5.00	< .01*	.56	.27
Social Networks/Blogs: Small - High	4.63	< .01*	.52	.23
Shopping Websites: Medium - High	4.55	< .01*	.51	.29
Social Networks/Blogs: Large - Moderate	4.30	< .01*	.48	.28
Social Networks/Blogs: Large - High	4.12	< .01*	.46	.26
Social Networks/Blogs: Medium - High	4.10	< .01*	.46	.25
Social Networks/Blogs: Medium - Low	4.07	< .01*	.45	.19
Social Networks/Blogs: Large - Low	3.82	< .01*	.42	.25
Shopping Websites: Large - High	3.72	< .01*	.41	.21
News Websites: Large - Moderate	3.41	< .01	.37	.20
Search Websites: Medium - High	3.19	.01	.34	.18
Shopping Websites: Medium - Low	2.94	.01	.30	.15
Search Websites: Large - Low	2.90	.01	.30	.22
News Websites: Large - Low	2.79	.01	.28	.18
Shopping Websites: Small - High	2.79	.01	.28	.18
Search Websites: Medium - Moderate	2.74	.01	.27	.20
News Websites: Medium - Low	2.67	.02	.26	.16
Shopping Websites: Large - Moderate	2.64	.02	.26	.17
Shopping Websites: Large - Low	2.30	.03	.21	.15
Search Websites: Medium - Low	2.15	.04	.19	.15
Shopping Websites: Small - Low	-2.10	.05	.18	-.13
Search Websites: Small - High	1.81	.09	.14	.13
Search Websites: Small - Moderate	1.77	.09	.14	.12
Shopping Websites: Medium - Moderate	1.75	.09	.13	.14
Shopping Websites: Small - Moderate	-1.75	.10	.13	-.09
Search Websites: Small - Low	1.71	.10	.13	.12
Search Websites: Large - Moderate	1.66	.11	.12	.14
News Websites: Medium - Moderate	1.61	.12	.11	.10

TABLE 53 (continued)

Comparison	*t*	*p*-value	r^2	MD
Search Websites: Large - High	1.02	.32	.05	.09
News Websites: Medium - High	.91	.37	.04	.07
News Websites: Small - High	-.83	.42	.03	-.05
News Websites: Large - High	.79	.44	.03	.07
News Websites: Small - Low	-.47	.64	.01	-.02
Social Networks/Blogs: Small - Moderate	.41	.69	.01	.03
Social Networks/Blogs: Small - Low	.24	.81	< .01	.02
News Websites: Small - Moderate	-.17	.87	< .01	-.01

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

Similarly to sensitivity, correlations with bias were conducted to determine if bias was related to previous web usage. Overall, no significant correlations were noted for any stimulus type or size and resolution combination. See Table 54 for correlations.

TABLE 54

CORRELATIONS BETWEEN BIAS AND WEBSITE USAGE

Size	Resolution	News Websites		Search Websites		Shopping Websites		Social Networks/Blogs	
		r	*p*	*r*	*p*	*r*	*p*	*r*	*p*
Small	High	-.28	.22	.37	.10	.07	.77	.06	.80
Small	Moderate	-.28	.23	.39	.08	-.08	.72	-.16	.50
Small	Low	.05	.83	-.12	.61	-.10	.67	.05	.83
Medium	High	-.46	.03	.44	.05	-.24	.29	-.16	.49
Medium	Moderate	.03	.91	.30	.18	.02	.94	.10	.66
Medium	Low	-.05	.86	-.01	.97	.08	.73	-.01	.97
Large	High	-.44	.05	.15	.52	-.07	.77	.12	.62
Large	Moderate	-.42	.06	.12	.60	-.36	.11	.01	.97
Large	Low	-.43	.05	.32	.16	-.28	.22	.36	.11

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

Weekly Internet usage was compared with the bias. Spearman correlations were conducted. The results demonstrated no significant relationships for Internet usage and bias.

APPENDIX L

STUDY 3 PLANNED COMPARISONS RESULTS TABLES

TABLE 55

STATISTICAL TESTS COMPARING SENSITIVITY FOR STIMULUS SIZES AT DIFFERENT RESOLUTIONS

Comparison	*t*	*p*-value	r^2	MD
Shopping Websites: Large - High	10.64	< .01*	.85	.40
Shopping Websites: Medium - Moderate	10.47	< .01*	.85	.37
News Websites: Large - High	10.32	< .01*	.84	.36
Shopping Websites: Small - High	10.05	< .01*	.84	.34
Shopping Websites: Large - Low	9.69	< .01*	.82	.36
Shopping Websites: Large - Moderate	9.59	< .01*	.82	.35
Search Websites: Large - Moderate	8.85	< .01*	.80	.38
Shopping Websites: Medium - Low	8.73	< .01*	.79	.31
Search Websites: Small - High	8.21	< .01*	.77	.35
Shopping Websites: Medium - High	8.07	< .01*	.77	.35
Social Networks/Blogs: Large - Moderate	8.05	< .01*	.76	.26
Search Websites: Small - Low	7.68	< .01*	.75	.28
Search Websites: Medium - Low	7.62	< .01*	.74	.34
Social Networks/Blogs: Large - High	7.61	< .01*	.74	.32
News Websites: Medium - High	7.42	< .01*	.73	.32
Search Websites: Large - Low	7.38	< .01*	.73	.35
Search Websites: Medium - Moderate	7.36	< .01*	.73	.33
Search Websites: Large - High	7.10	< .01*	.72	.39
News Websites: Medium - Low	6.95	< .01*	.71	.26
Search Websites: Small - Moderate	6.79	< .01*	.70	.31
News Websites: Large - Moderate	6.40	< .01*	.67	.29
Search Websites: Medium - High	6.39	< .01*	.67	.33
News Websites: Large - Low	6.39	< .01*	.67	.28
Social Networks/Blogs: Large - Low	6.31	< .01*	.67	.27
Social Networks/Blogs: Medium - High	6.25	< .01*	.66	.26
Social Networks/Blogs: Medium - Moderate	5.76	< .01*	.62	.28

TABLE 55 (continued)

Comparison	*t*	*p*-value	r^2	MD
Social Networks/Blogs: Small - High	5.38	< .01*	.59	.24
News Websites: Medium - Moderate	4.45	< .01*	.50	.24
Social Networks/Blogs: Medium - Low	4.02	< .01*	.45	.18
Shopping Websites: Small - Moderate	2.88	.01	.29	.17
Shopping Websites: Small - Low	2.49	.02	.24	.10
News Websites: Small - High	2.25	.04	.20	.11
News Websites: Small - Moderate	2.17	.04	.19	.10
Social Networks/Blogs: Small - Moderate	1.96	.06	.16	.08
Social Networks/Blogs: Small - Low	1.73	.10	.13	.08
News Websites: Small - Low	1.47	.16	.10	.05

Note: * indicates a significant comparison. Mean differences are differences in backtransformed means. Holm-Bonferroni adjustments were applied to control for familywise Type I error.

TABLE 56

STATISTICAL TESTS COMPARING BIAS B''_D FOR STIMULUS SIZES AT DIFFERENT RESOLUTIONS

Comparison	*t*	*p*-value	r^2	MD
Social Networks/Blogs: Large - Moderate	6.85	< .01*	.70	.16
Social Networks/Blogs: Medium - Moderate	6.22	< .01*	.66	.13
Shopping Websites: Medium - High	5.51	< .01*	.60	.12
Social Networks/Blogs: Large - High	5.39	< .01*	.59	.13
Social Networks/Blogs: Medium - High	5.06	< .01*	.56	.12
Social Networks/Blogs: Large - Low	5.03	< .01*	.56	.13
News Websites: Large - Moderate	4.64	< .01*	.52	.11
Social Networks/Blogs: Medium - Low	4.29	< .01*	.48	.11
News Websites: Medium - Low	3.81	< .01*	.42	.11
Shopping Websites: Large - Moderate	3.78	< .01*	.42	.10
Shopping Websites: Large - High	3.45	< .01	.37	.08
News Websites: Large - Low	3.36	< .01	.36	.10

TABLE 56 (continued)

Search Websites: Large - Low	3.35	< .01	.36	.09
Shopping Websites: Small - Low	-2.97	.01	.31	-.11
Search Websites: Medium - High	2.69	.01	.27	.07
News Websites: Medium - Moderate	2.68	.01	.26	.08
Search Websites: Medium - Moderate	2.65	.02	.26	.08
Shopping Websites: Large - Low	2.53	.02	.24	.07
Social Networks/Blogs: Small - High	2.49	.02	.24	.09
Search Websites: Small - Low	2.42	.03	.23	.07
Shopping Websites: Small - Moderate	-2.30	.03	.21	-.07
Search Websites: Medium - Low	2.24	.04	.20	.06
Shopping Websites: Small - High	2.19	.04	.19	.06
Shopping Websites: Medium - Low	1.95	.07	.16	.05
Search Websites: Small - High	1.59	.13	.11	.05
Search Websites: Large - Moderate	1.55	.14	.11	.05
Shopping Websites: Medium - Moderate	1.52	.14	.10	.04
News Websites: Large - High	1.38	.18	.09	.05
Search Websites: Small - Moderate	1.36	.19	.08	.04
Search Websites: Large - High	1.00	.33	.05	.03
News Websites: Medium - High	.94	.36	.04	.03
Social Networks/Blogs: Small - Moderate	.81	.43	.03	.03
Social Networks/Blogs: Small - Low	-.60	.56	.02	-.02
News Websites: Small - Low	.60	.56	.02	.02
News Websites: Small - Moderate	-.41	.69	.01	-.01
News Websites: Small - High	-.12	.90	< .01	< -.01

Note: * indicates a significant comparison. Holm-Bonferroni adjustments were applied to control for familywise Type I error.