LUNDS UNIVERSITET

DEPARTMENT OF PSYCHOLOGY

The Effect of Color in Website Design: Searching for Medical Information Online

Gabriel Nordeborn

Magister thesis spring 2013

Supervisor: Geoffrey Patching

THE EFFECT OF COLOR IN WEBSITE DESIGN

Abstract

As the spread of the internet continues, more people than ever search for medical information online when in need of health advice. Color is a prominent part of website aesthetics, and has been shown in the literature to have both positive and negative psychological effects. A total of 120 participants took part in the present study where the effect of the colors grey, blue, orange and red was tested on the task of searching for medical information online. An experimental website was manipulated with different color schemes, inspired by existing popular websites with the purpose of providing medical information online. Participants then used the website to find and comprehend information about a fictitious disease. Overall, the results showed no statistically significant effect of website color on either ratings or time spent using the website. A three-way interaction between color, gender and ratings was found, where females rated the red website significantly higher than males on almost all dimensions. The results are discussed in light of similar findings of the color red and its negative effect on males. Future research involving other areas where the potential effect of color might be of benefit is suggested.

Keywords: *human-computer interaction*, *psychology, medical information, color, usability*

The Effect of Color in Website Design

Color is regularly used to guide behavior, such as green traffic lights, and please our senses. Color is something that affects us every day, and is ever present in our lives. Since Cohn's (1894) pioneering research on color, lighting and emotions, various notions of color perception, preference, and aesthetics have been investigated widely (see Whitfield & Wiltshire, 1990, for a review). However, despite considerable research interest in color, there is almost a complete absence of studies examining potential applications of color for website design. Here, the role of color for website design is investigated in the context of searching for medical information using the internet by manipulating the color of an experimental website.

Medical ailments are at times present in our own and our significant others' lives. Yet, with the rise of internet as something available to most people in the industrialized world, there has never before been as much information as easily available to us as now. This is assured by the introduction of computers in most people's lives, but also with the addition of smart phones, tablets and wireless internet connectivity technology like 3G, and more. Given this shift in technology, where more than 60% of the population of Europe has access to the internet via technology like computers (http://www.internetworldstats.com/stats4.htm), scientific communities have developed to investigate humans' interaction with computers and machines. The most prominent field of research in this regard is called Human-Computer Interaction (HCI). HCI is a multidisciplinary research area focused on the interaction between humans and technology, which draws on many different disciplines, such as Computer Science, Ergonomics, and Psychology.

A major part of HCI concerns "usability", which refers to bridging the gap between designers, developers and the actual individuals using the technology developed. For instance, Donald Norman in his book "The design of everyday things" (Norman, 1988), strongly impacted the field of HCI, by focusing on the gaps that frequently exist between the design of the technology and the skills and knowledge of the intended users. Here, usability is a simple concept; things should be usable by the people that are intended to use them. Yet, it is a complex research area, with countless factors potentially affecting the usability of a product. The role of color in the design of websites for searching medical information is of primary focus in the present study.

Medical Information Online

Medical information online should be reliable, accurate, and accessible. Since the internet was introduced to a widespread audience, access to medical information has become readily available, especially with the recent inflation of mobile devices and their access to the internet. Medical information online was present early in the history of internet, and, today, many people use the internet to seek medical information when required.

Studies indicate that medical information was an early feature of general interest on the internet (O'Connor & Johansson, 2000; Diaz, Griffith, Reinert, Freidmann, & Moulton, 2002; Eysenbach & Köhler, 2002). In a study conducted in a gastroenterology clinical setting by O'Connor and Johanson (2000), it was found that approximately half of the patients in the study, with access to the internet, had used the internet for finding medical information. The majority of users also stated that they intended to continue using the internet as a source of medical information in the future. The study also revealed that 96% of the participating patients had been recommended to use the internet to find medical information by their physician (O'Connor & Johanson, 2000). Diaz et al. (2002) found similar indications in their study of patients' use of the internet for finding medical information, where more than half of the participants sought medical information online in addition to the information provided by their physicians. In the study by Diaz et al. (2002), 60% of the participants rated the information they found online as “as valuable as, or more valuable than” the information they received from their physicians. This indicates the prized value of medical information online, even in its early history.

A study by Eysenbach and Köhler (2002) investigated the actual procedure people used to find medical information online. First, participants made comprehensive lists of what they used to judge the credibility of a website with medical information online. However, when their actual behavior and strategies were investigated in post-experiment interviews, very few had actually verified the factors they stated as being important in finding credible information (Eysenbach & Köhler, 2002). The participants stated that factors like a professional design, a scientific or official “touch” and ease-of-use were important for judging the credibility of the site providing the information. Nevertheless, when asked after performing the experimental task of actually searching for medical information online, very few had looked at the “About us”-pages, or even remembered on what page they had successfully found the information

(Eysenbach & Köhler, 2002). The indication is, therefore, that there could be two or more relevant dimensions when searching for medical information online: I) the dimension of appearances that the users themselves claim they are using as a basis for their judgments of the website, II) the dimensions of pure functionality and performance; how well the user could actually use the website to find and comprehend the required information.

Today, at least in the US, more than 70% of citizens with access to the internet report that they have sought medical information online in the previous year (Kuehn, 2013). Other notable statistics are: (a) eight out of 10 online health inquiries begin by searching using a search engine, like Google (www.google.com), (b) half of the health information searches are reportedly on behalf of someone else, and (c) females are more prone than males to search for medical information online (Kuehn, 2013).

Reportedly, in Canada the internet is the third main source of information for medical information, second only to actual physicians and newspapers (Krewski et al., 2006). Also, there may be a relation between consuming medical information online and the amount of face-to-face health care required. Searching for medical information online may lower the amount of face-to-face health care required, and in that way help reduce the work load of health care professions, freeing up a valuable resource for those in most need (Suziedelyte, 2012). Yet, as the recent swine flu alarm in 2009 revealed, it is vital to have access to accurate medical information quickly when required. Hilton and Smith (2010) found that during the 'swine flu' (H1N1 influenza virus) epidemic in 2009, with the large amounts of sometimes conflicting information available through all sorts of sources, certain groups of people perceived greater risks to their health from the vaccine than the actual disease itself. Moreover, Gilmour, Huntington, Broadbent, Strong, and Hawkins (2012) found that a lot of nurses' work consisted of correcting misinterpreted medical information found by their patients online. A task that Gilmour et al. (2012) found was often not completed optimally, potentially leaving patients with misinterpreted medical information about their physical conditions. Making sure that all of the available medical information online is reliable and correct is an impossible task due to the nature of the internet. However, steps may be taken to ensure that the sources that do provide correct information do so in such a way that attracts those people searching for medical information to reliable websites.

Attitude Affects the Use of Systems and Interfaces

Research has shown that positive attitudes towards systems and interfaces can be formed from only perceived efficiency and usability, without that efficiency and usability necessarily existing for real (Schmidt, Liu, & Sridharan, 2009). This means that a system that is perceived as efficient and usable may in reality actually not be neither efficient nor usable. Systems for providing medical information online is no different, and as people will use these systems to find medical information, often also for their significant others (Rice, 2006; Kuehn, 2013), it is important that these systems provide good, real usability. This to not only get perceived as usable, but actually be designed to perform tasks as efficiently as possible.

Attitude has long been established as a good determinant of actual system use (Davis, Bagozzi, & Warshaw, 1989; Ajzen, 1989). Davis et al. (1989) proposed that attitude, together with subjective norms, and perceived usability, were good determinants of whether a system gets used or not. Ajzen (1989) has also noted that intention to use a system or not is tied to the user's attitude toward that system, and Davis et al. (1989) found that intention to use a system is highly related to the likelihood of actually using that system.

Further expanding on usability and its importance in systems and the likelihood of system use, Schackel (1991) discusses a number of factors affecting whether a system is accepted for use or not. Schackel (1991) argues that this process is dependent on 3 dimensions; (a) utility, if the system can do what is needed of it, (b) usability, if the system can actually be used for what it is supposed to be used for, and (c) likability, if you like the system or not. According to Schackel (1991), a combination of high compatibility with these 3 dimensions leads to system acceptance.

In summary, there is a growing body of evidence to suggest a relation between attitude, satisfaction and usability. Compounded with the ambiguous nature of attitude where actual effectiveness is not necessary to attain good perceived usability, it is therefore necessary to assess the performance of a system along two main dimensions; (1) the perceived usability and general experience of using the system, and (2) the actual effectiveness of the system.

Color

When investigating the literature addressing color and its potential effects, it is evident that color preference is of widespread concern (cf. Whitfield & Wiltshire, 1990). Since Cohn (1894), preferences have been investigated in many different contexts, ranging from creating

guidelines for how color should be used in buildings (O'Connor, Whitfield, & Wiltshire, 1989) to how color affects the perception of potato chips and their freshness (Maga, 1973). Early research was mostly conducted with single and pairs of colors, with color preference generally investigated in the terms of hue (cf. Whitfield & Wiltshire, 1990). Note, color consists of three dimensions; (a) hue, which controls the perceived qualitative experience of color like red, green or blue, (b) chroma, which controls the saturation of the color, and (c) luminance, which controls the perceived brightness (Gorn, Chattopadhyay, Sengupta & Tripathi, 2004).

Color in psychology. In the psychological literature, colors are known to have various emotional and psychological properties (Lichtle, 2007). Blue has traditionally been a color associated with positive emotional effects, such as relaxation (Gorn et al., 2004), pleasant feelings (Valdez & Mehrabian, 1994) and increased likelihood for purchasing products (Bellizzi & Hite, 1992; Crowley, 1993; Babin, Hardesty, & Suter, 2003). Blue is generally an appreciated color in many different cultures (Marcus & Gould, 2000). Lichtle (2007) describes blue as a color generally associated with security, wealth and trust.

As compared to blue, red is considered a color with fewer positive properties (Bellizzi & Hite, 1992). Where blue supposedly increases purchase intention, red is not effective (Bellizzi & Hite, 1992; Crowley, 1993). Elliot, Maier, Moller, Friedman, and Meinhardt (2007) found that the mere presence of the color red could impair performance attainment, which they argue could be due to red being associated with failure. In their series of experiments, participants were cued with the color red before doing a performance-related task, which caused the participants to perform worse than their control group equivalents (Elliot et al., 2007). Gnambs, Appel, and Batinic (2010) found similar but gender-dependent effects, where males performed worse than females when exposed to a red, as opposed to a green progress bar when doing a performance-related online task. According to Gnambs et al. (2010) this is potentially related to a gender-dependent stereotype threat. Further, building on the research of the negative effect of red, Lichtenfeld, Maier, Elliot, and Pekrun (2009) found that merely reading the word red, without any actual color being displayed, affected the performance of participants negatively.

Orange has been found to have both positive (Bonnardel, Piolat, & Le Bigot, 2011) and negative (Babin, Hardesty, & Suter, 2003) properties. Babin et al. (2003) found that orange decreased shopping intention in a retail store context, where blue increased the shopping

intention. This effect was however nullified as the lighting was decreased (Babin et al., 2003). In the context of shopping, Lichtle (2007) found orange to be associated with “cheapness”. Alternatively however, Bonnardel et al. (2011) found orange and blue to be among the most preferred colors with their group of participants. In sum, there is no consensus about the psychological effects of orange in the relevant literature.

Grey has usually been considered a neutral color (Bonnardel et al., 2011). Grey has been used as the neutral baseline in various studies investigating colors' psychological effect (Elliot et al., 2007; Lichtenfeld et al., 2009; Elliot & Aarts, 2011). Generally speaking, grey is used as a baseline because no specific psychological effects are thought to be aroused from being exposed to this particular color.

Notably, Bonnardel et al. (2011) found that the colors of blue and orange helped facilitate the comprehension of material available inside their experimental website. This has given reason to investigate color potentially increasing website usability, as examined in the present study.

Website Design

The literature indicates a strong relation between satisfaction, perceived usability, aesthetic design, and factors like credibility and trustworthiness, in determining the use of websites. For example, Robins and Holmes (2008) found that a high aesthetic treatment yields a higher perceived credibility, a process which Robins and Holmes' (2008) study indicates happens in mere seconds. Likewise, Schenkman, and Jönsson (2000) found aesthetics to be relevant in website design, as their study suggested beauty to be one of the biggest factors in how preferable websites were perceived. Flavián, Guinalíu, and Gurrea (2006) found that both satisfaction and the perceived usability of websites were related to trustworthiness. In similar vein, Cyr, Kindra, and Dash (2008) found a relation between visual design and trustworthiness.

Notably, the process of forming opinions about website seems to happen quickly. Research shows that it takes people little time to form stable opinions about websites, generally approximately 500ms (Tractinsky, Cokhavi, & Kirschenbaum, 2004), but some even as quick as 50ms (Tractinsky, Cokhavi, Kirschenbaum, & Sharfi, 2006).

Bringing Together Color and Medical Information Online

Color and its potential effects have been said to be context dependent, and sensitive to being investigated in a natural context (Bonnardel, Piolat, & Le Bigot, 2011). Given the background to both color as a psychological subject and its role in HCI, combined with rapidly increasing use of the internet to find medical information, it is particularly timely to investigate colors' potential effect in the context of searching for medical information online.

One way to examine this issue is to attempt to mimic how people search for medical information online, and then systematically manipulate the color of the website the individual uses to find the relevant information. In the present study, this question was addressed by constructing an experimental website, and by systematically manipulating the color of the website. After Bonnardel et al. (2011), the website used in the present study was designed to resemble already existent websites providing medical information online, especially in regards to the usage of color. The manipulation of the color of the experimental website was designed to be sufficiently comprehensive to give the website a strong color identity, but not divert in any crucial respect from already existent websites in its use of colors. To ensure that participants in the experiment actually interacted with the experimental website, an experimental task of answering a multiple-choice question was employed. This experimental question was based on a fictitious disease, which the participant was required to answer using information available by searching the website.

The Website

Technically, the experiment consisted of three parts; (a) the website itself, (b) customized code to control the manipulation of the website, register relevant data and ensure that the website was accessible only to those who had consented to take part in the study, and (c) a post-experiment questionnaire.

The website itself was constructed using Wordpress (http://www.wordpress.org). Wordpress is a free and open source content management system for designing websites. Manipulation of the coloring was made using CSS (Cascade Style Sheet). CSS is regularly used in aesthetic design of websites.

In an attempt to mimic existing sites with similar purpose, the website design drew upon five real, existing, medical informational sites. These websites were found by way of The Medical Library Association of America (http://www.mlanet.org/resources/userguide.html).

An initial pilot investigation of the selected websites revealed that the most common factors for the design of the websites was (a) 60% had a solid background color, (b) 80% had a proper logotype and search bar, and (c) 60% provided easy access to a menu listing diseases alphabetically by letter, from A to Z. Therefore, design of the present experimental website incorporated a solid background color, a logotype in the header of the page, a search bar in the right most column of the page, and easy access to a menu listing diseases from A to Z.

The information contained in the experimental website. In order to preserve credibility and trustworthiness, the website itself contained a considerable amount of medical information over and above the information about the fictitious disease related to the experimental question. This additional medical information was included because the scales of interest in the post-experiment questionnaire were aimed at garnering a judgment of the experience of using the website itself, and not exclusively navigating to the page required to answer the experimental question. Therefore, navigation on the website, outside of completing the experimental task, needed to be realistic as well. Six to eight entries of random diseases derived from Wikipedia (www.wikipedia.org) were entered for each letter in the alphabet. Additionally, eight diseases randomly picked from the numbers 0 to 9 were also included. This resulted in, excluding the fictitious disease, a total of 237 entries about random diseases being added to the experimental website.

The entries derived from Wikipedia were modified before being incorporated in the present website. First, all external hyperlinks were removed from the entries, to ensure no one navigated away from the site via external links during the experiment. Second, tables of contents, suggested extra readings, and repetitive parts of the entries were removed. Third, all references were removed for coherence with the text of the fictitious disease.

Customized code to control the experiment. The technical mechanisms of the experimental procedure itself were controlled using customized code in PHP: Hypertext Preprocessor (http://www.php.org). A MySQL database (http://www.mysql.org) controlled by PHP was used for data registration and storage.

Each participant was required to give their consent to partake in the study in order to be granted access to the experimental interface, and each participant could not access the post-experiment questionnaire until the experimental task had been successfully completed. This was to ensure that it was impossible to skip to the end questionnaire without using the interface, and so limit false entries from being registered.

The design of the experimental website. The website was called "NetDoc", which was presented as an online service for easily accessible medical information. The interface contained a logotype at the top, consisting of the name "NetDoc" and a symbol in the shape of a cross. The color of the symbol varied with the color schemes. It also contained a menu ranging from A to Z, giving access to the various articles about diseases available in the interface by the first letter of the disease name. This A to Z menu was located under the logotype at the top of the page. There were two columns under the A to Z menu, with the leftmost displaying the actual content of the current page, and the rightmost containing a search function and several hyperlinks to various pages. Finally, at the bottom of the website, a small text was displayed to remind each participant of the experimental nature of the website, along with brief instructions about how to view the experimental question.

Variables and Measures

As discussed, providing medical information online is a two-fold task. First, the website providing the information needs to be perceived as beneficial on a number of factors such as design, and efficiency. Second, the website needs to be as efficient as possible in real, and not just perceived, performance. Therefore, the variables assessed in the present study aimed to measure both perceived and actual usability in order to provide comprehensive data for analyzing the potential effect of color in websites providing medical information online.

The scales used in the present study were derived from a study by Teo, Oh, Liu, and Wei (2003), who investigated the effect of interactivity in an experimental website on a number of dimensions related to the experience of using the website. More specifically, Teo et al. (2003) based their questionnaire on the dimensions of effectiveness, efficiency, value, satisfaction and attitude. Their questionnaire was based on earlier research concerning website usability and experience (Teo et al., 2003). The scales were all shown to be valid and reliable in the original article (all Cronbach alpha's > 0.8), and the convergent validity of the scales was proven via three convergent validity tests (Teo et al., 2003). In addition, a satisfactory discriminant analysis was also performed by the authors, which suggests that the various questions employed are independently measuring the separate dimensions they refer to (Teo et al., 2003). However, the original scales of Teo et al. (2003) were constructed with the aim of assessing the effect of interactivity as their main goal. Consequently, certain modification of some of the questions was required for present purpose. An example of this alteration is in the

scale of efficiency, where “The Web site allowed me to make a decision quickly” was changed to “The Web site allowed me to answer the question quickly”. The scale of effectiveness was excluded in the present study, because it was aimed exclusively at measuring interactivity in the original article (Teo et al., 2003).

Measure of effectiveness. In the present study, the effectiveness of using the website was assessed by recording the time taken from reaching the correct page containing the information required to answer the experimental question, to when the experimental question was answered correctly. This measure allowed for the registered time to reflect only the time taken reading and making sense of the relevant information. An aim was to make sure the participant had understood the experimental task, which is indicated by them looking for and finding the correct page. This measure will be referred to as search time in analysis.

The Color Manipulation

The different base colors for the four color schemes used in the manipulation of the color of the website were all derived from a previous study by Bonnardel, Piolat, & Le Bigot (2011). Their study showed a number of effects of colors, ranging from some colors being more preferable than others, to some colors enhancing the comprehension of website material (Bonnardel et al., 2011). In particular, Bonnardel et al. (2011) found a positive effect of both orange and blue on comprehension of their website material. This gave reason to believe there might be positive effects on performance in the present study of using the colors of orange and blue.

Every effort was made to provide an identical experience for each participant in terms of the chroma and luminance of the colors displayed. However, there is no reliable way of fully ensuring that every participant had the same color experience, as local screen settings inevitably vary. Nonetheless, the experimental manipulation was made only by altering the hue.

The experiment used four color manipulations: Grey, blue, orange, and red.

1. Grey was chosen due to its aforementioned quality of being perceived as a neutral color (Bonnardel et al., 2011), and because it has been frequently used as a baseline in previous color experiments (Elliot et al., 2007; Lichtenfeld et al., 2009; Elliot & Aarts, 2011)
2. Blue has consistently been deemed as a color with positive properties (Gorn,

Chattopadhyay, Sengupta, and Tripathi, 2004; Lichtle, 2007; Marcus & Gould, 2000), and used widely throughout the literature, for example in shopping contexts (Bellizzi & Hite, 1992; Crowley, 1993; Babin, Hardesty, & Suter, 2003). Blue has also been linked to increased performance of comprehending material in websites (Bonnardel et al., 2011).

3. Orange was chosen as there are conflicting findings about the usefulness of orange in the literature, with some claiming positive effects (Bonnardel et al., 2011), and others claiming negative effects (Babin, Hardesty, & Suter, 2003). Orange, like blue, has also been linked to increased performance of website comprehension (Bonnardel et al., 2011).
4. Red has traditionally been considered a color with bad properties (Bellizzi & Hite, 1992), with the potential to impair performance and decreasing positive experience (Elliot et al., 2007).

Hypotheses

On the basis of previous research in the area the following hypotheses inspired the present study:

Hypothesis 1: *There will be differences between groups on the ratings of the scales of satisfaction, efficiency, attitude and value, as measured by the post-experiment questionnaire depending on the color of the experimental website. There may also be differences depending on gender among the groups.*

Hypothesis 2: *There will be a difference between the groups, as randomly assigned to the different color websites, in the time taken from finding the correct informational page to answering the experimental question correctly. There may also be a difference depending on the gender of participants.*

Method

Participants

120 participants took part in the experimental study. Participants were recruited by way of

advertisement on various electronic social media sites, such as Facebook (http://www.facebook.com). A total of 19 participants were excluded from the analyses; 13 because they failed to complete the study questionnaire; 3 because they stated they were colorblind; 1 because they failed to specify being either female or male; and 2 because the amount of incorrect tries in answering the experimental question in combination with the time taken to answer the question gave strong reason to believe the participants had not completed the experiment in a valid way. The final sample consisted of 101 participants; 49 female and 52 males. On starting the experiment, each participant was randomly assigned to one of four experimental groups. A Chi-square test revealed no significant differences in number of females and males in each group, $\chi^2(3, N = 101) = 6.12, p > .05$; and so the randomization worked successfully.

Experimental Design

The experiment utilized a between-participants design with four experimental groups. The manipulation in the experiment was the color scheme of the website. This varied over four different colors; grey, blue, orange, and red. Except for the color varying between groups, the participants all used identical websites in the experiment.

The experimental task consisted of answering a multiple-choice question about a fictitious disease named “Linearian Disease”. Each participant had to navigate the website to find the correct page containing the information required to answer the question. The participant then needed to read and comprehend the information to answer the question correctly. Upon finding the correct information and answering the question successfully, the participant was directed to a post-experiment questionnaire.

Manipulation of the color. The parts of the experimental website that were systematically manipulated by color scheme were: (a) background color, where the background color of the website was fully manipulated, (b) the logotype color, with the symbol in the logotype being manipulated, (c) the color of the headlines on the website, with all of the headlines being manipulated, and finally (d) the hyperlinks in the navigation menu, where the color of the hyperlinks changed when hovering over them. All else on the website was held constant for every participant. All manipulations, except (d), the hovering of the hyperlinks, can be viewed in Figure 1 and Figure 2 below. Figure 1 and Figure 2 show the difference between the blue

and the red manipulations of the experimental website.

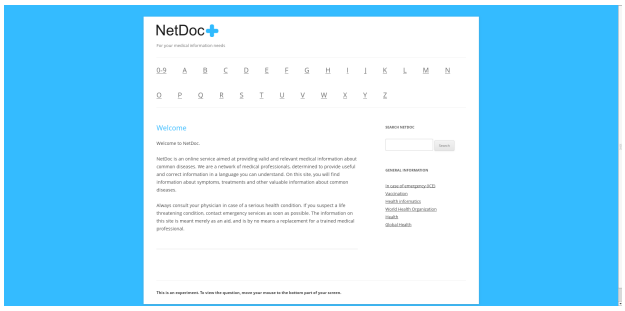

Figure 1. The blue color scheme of the experimental website.

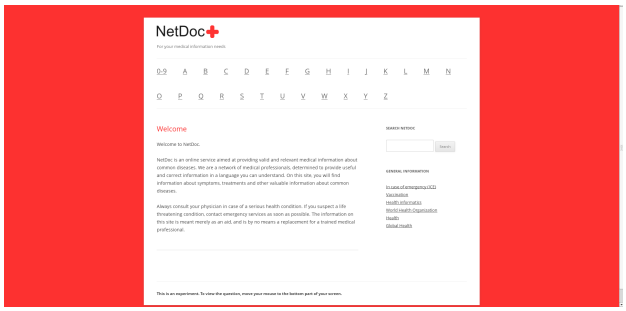

Figure 2. The red color scheme of the experimental website.

Post-experiment questionnaire. After successfully completing the experimental task, the participant was presented with a questionnaire assessing the various dimensions of interest, demographical data and also an open question where any other comments could be made. The dimensions assessing the experience of using the website consisted of; (a) satisfaction of using the website, (b) perceived efficiency of the website, (c) perceived value of the website, and (d) attitude towards the website.

Satisfaction and efficiency were measured with four questions each, where the participant was asked to rate how much he or she agreed with a statement regarding the experience of using the website. For example, (a) satisfaction "I feel very satisfied with my visit to the Web site", and (b) efficiency "I was able to access the information I needed very quickly". Participants were requested to rate each statement using a Likert scale ranging from 1 to 7, where 1 represented strongly disagree, 4 neutral and 7 strongly agree.

Attitude and value were also assessed by way of Likert scales, ranging from 1 to 7. Attitude contained three items, and value contained four items. In this case however, participants were requested to indicate the extent to which they agreed with adjectives describing the website. Example items from (c) attitude are “worthless”, represented by the value of 1, to “valuable”, represented by the value of 7. From (d) value, “dislike”, represented by the value of 1, to “like”, represented by the value of 7.

Experimental Task

The experimental task was designed to ensure participants interacted with the website in a manner similar to everyday searching for medical information online. In this respect the

experimental task was constructed as a multiple choice question. All participants were instructed to find the answer to the multiple choice question within the website. The question itself was available to each participant at all times via a pop-up window located at the bottom of the screen. To ensure the question did not interfere with participants' experience of using the website, the question was hidden by default. This meant that the participant had to move their cursor to the bottom 20% of the screen to show the question, which then popped up for as long as the cursor remained in the bottom 20% of the screen.

The experimental question. The experimental question was "Which of the following descriptions characterizes symptoms of Linearian Disease?", with the following choices available: (a) "It always produces nausea and vomiting for children.", (b) "Roughly a third of the adults gets a headache and a sore throat.", (c) "In certain cases the sickness might last a full week.", and (d) "The virus cannot be inactivated."; (c) was the correct choice. The question and alternatives remained the same for every participant, and the order in which the choices were displayed was randomized.

The fictitious disease and its informational text. The text which described the fictitious disease "Linearian Disease" used in the experimental task was based on Wikipedia's text describing common influenza (http://en.wikipedia.org/wiki/Influenza). The text was modified so as to not resemble influenza, but something similar. The use of a fictitious disease, instead of a real one, was necessary in order to ensure that no one (a) knew the answer to the question and so did not need to interact with the website, and (b) had a personal experience of the disease, potentially influencing their experience of using the website. However, as the goal of the study was to keep the experiment as realistic as possible by mimicking how people actually go about searching for medical information online, the fictitious text was initially screened by a qualified medic at the Karolinska Institute, Stockholm, Sweden. This was to ensure there were no medical contradictions in the text.

The text, which in size was approximately equal to a regular A4 size page, consisted of four paragraphs; introduction, symptoms, causes, and treatments. The alternatives for the question were constructed so that each alternative was relevant to the text, but only one alternative was correct. The correct alternative was located in the third paragraph, and the way it was phrased ensured that it could not be found merely by using the built in search function of the website, nor the search function available in any modern web browser. These measures were taken in an attempt to ensure that participants read and comprehend enough of the text to

be able to answer the question correctly.

Pilot Study

A pilot study was initially conducted to assess the experimental website design, along with the scales and dimensions used in the post-experiment questionnaire. 5 psychology students aged from 21 to 28 ($M = 24.40$, $SD = 2.70$) took part in the pilot study. In the pilot study, the participants were instructed to iterate through the experiment twice; the first time as if they were actually taking part in the experiment, and the second time to assess the experimental website in more detail. An aim was to reveal any issues in the experiment that might need addressing. On the basis of the pilot study, a number of the post-experiment questionnaire items were adjusted in their phrasing to avoid extreme scores. For example, in order to increase the range of the possible responses the question “I feel satisfied with my visit to the Web site” was modified to “I feel very satisfied with my visit to the Web site”. In addition, a number of factors related to the function of the website, such as underlining certain hyper links, were adjusted.

Results

Data analysis was conducted in three steps. First, correlations among participants' ratings of Satisfaction, Efficiency, Value, Attitude and the measure of time spent, measured in seconds, answering the question (search time) were investigated to explore relations between search time and the rating variables.

Second, participants' search time was investigated by experimental group (website color) and gender. These data were submitted to a univariate analysis of variance (ANOVA) with search time as the dependent variable, along with the two between-participants factors, website color (grey, blue, orange, and red) and gender (female, male). The aim here was to shed light on the influence of the color of the experimental website on the efficiency of participants’ information search.

Third, participants’ ratings of Satisfaction, Efficiency, Value, and Attitude, as measured in the post-experiment questionnaire were examined by website color. All ratings were converted to standardized *Z*-scores prior to analysis, as one of the ratings (Attitude) contained less items than the others. These data were submitted to a mixed design ANOVA with two between participants’ factors; website color (grey, blue, orange, and red) and gender (female,

male), and one within-participant factor ratings (Satisfaction, Efficiency, Value, and Attitude). Search time variable was added as a covariate, to partial out differences between men and women in the time spent answering the experimental question. Mauchly's test showed that the assumption of sphericity was violated, $\chi^2(5) = 13.56$, $p < .05$, so degrees of freedom were corrected using Greenhouse-Geisser estimates (Greenhouse & Geisser, 1959) of sphericity ($\varepsilon = 0.98$).

Correlations among Variables

Table 1 shows relations between the variables as examined using Pearson's r. This analysis failed to reveal any statistically reliable relations between the rating variables and search time. Initial examining of scatterplots failed to reveal any curvilinear relations between these variables and further examination of the data using Spearman's correlation revealed a similar pattern of results.

Table 1. Pearson Correlations among variables of ratings and the variable of search time.

Measure	Satisfaction	Efficiency	Value	Attitude	Search time
Satisfaction	1	0.56**	0.47**	0.62**	0.15
Efficiency		1	0.33**	0.41**	0.06
Value			1	0.53**	0.06
Attitude				1	0.09

Notes. $N = 101$. ** = $p < .01$.

Search Time

The univariate ANOVA assessing search time in relation to website color and gender revealed a statistically significant main effect of gender, $F(1, 101) = 4.165$, $MSE = 16282.61$, $p < .05$, partial $\eta^2 = 0.040$, where males spent significantly less time ($M = 116.46$, $SD = 48.64$) searching for information than females ($M = 145.63$, $SD = 73.93$). No statistically significant main effect of website color was found, $F(3, 101) = 0.648$, $MSE = 2534.93$, $p > .05$, partial $\eta^2 = 0.020$; and no statistically reliable interaction effect was found between website color and gender, $F(3, 101) = 1.120$, $MSE = 4379.19$, $p > .05$, partial $\eta^2 = 0.035$.

Ratings

Analysis of participants ratings of the website by website color using a mixed design ANOVA, with search time used as covariate, revealed a statistically significant interaction effect between website color, gender and ratings, $F(8.22, 251.93) = 2.06$, $MSE = 1.097$, $p < .05$, partial $\eta^2 = 0.060$. Figure 1a shows the mean average of ratings (transformed to *Z*-scores for comparative purpose) of the websites made by males, and Figure 1b shows the mean ratings (*Z*-scores) made by females. No other statistically significant main effects or interactions was obtained (all *p*s > .05).

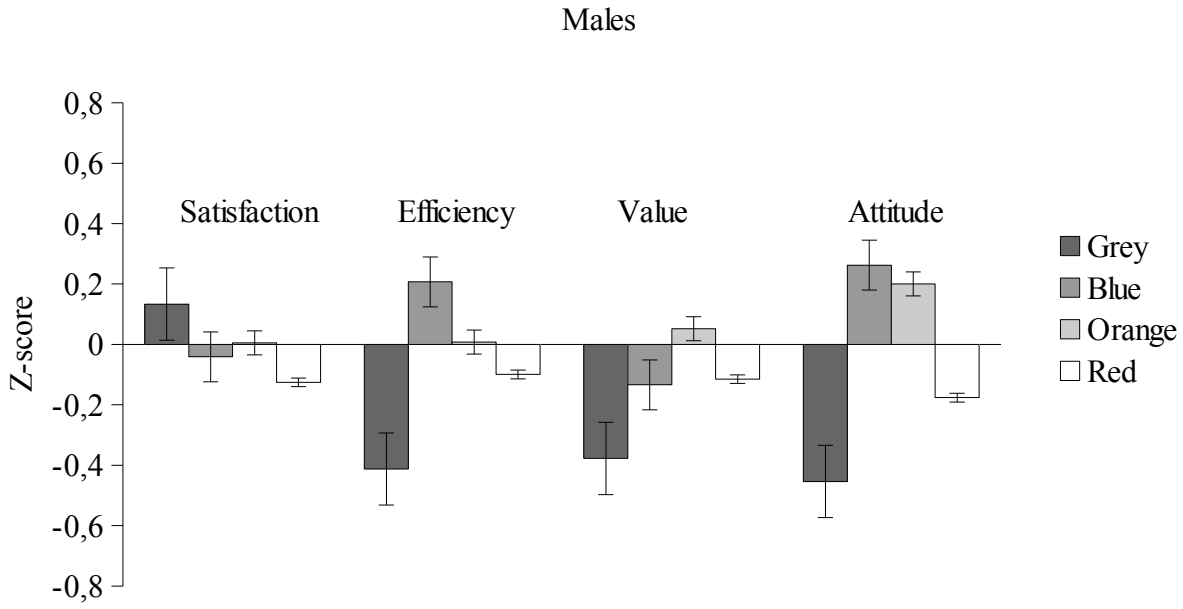

Figure 1a. Ratings of the websites by males. Standard error bars show the standard error of the mean.

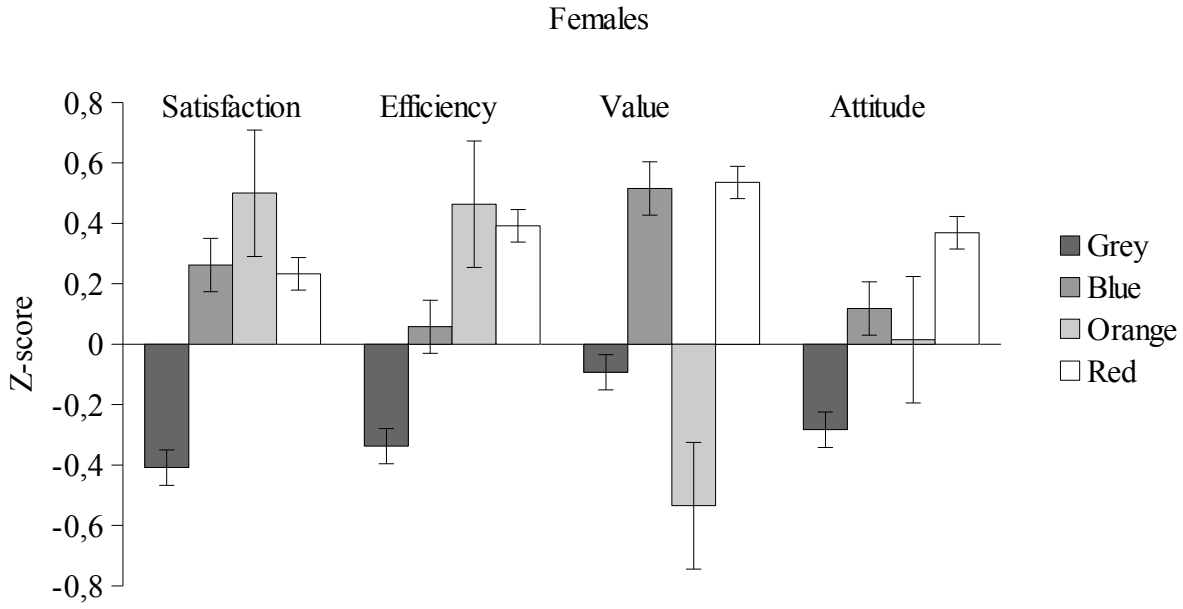

Figure 1b. Ratings of the websites by females. Standard error bars show the standard error of the mean.

Post hoc analyses of the interaction between color, gender and rating was examined further using Tukey HSD tests. This analysis revealed that females rated the red website significantly higher than males ($p < .01$) on Efficiency (females $M = 0.535$, males $M = -0.115$), Value (females $M = 0.364$, males $M = -0.176$) and Attitude (females $M = 0.388$, males $M = -0.099$). No other coherent pattern of statistically significant gender differences in website ratings was found.

Discussion

This study aimed to investigate the effect of color on the ratings and performance of using an experimental website to search for medical information online. This was done by manipulating the color of an experimental website, and randomly assigning participants to the different colored websites. The results failed to reveal statistically significant effects of color on the ratings of the website, and found no statistically reliable difference in participants search time when using the website. However, a statistically significant interaction was found between color, gender and ratings, when using search time as a covariate. Post hoc analyses using Tukey HSD tests revealed that females rated the red website significantly higher than males did on the dimensions of efficiency, value and attitude.

Color has previously been linked through aesthetics to both the credibility (Robins & Holmes, 2008) and trustworthiness (Cyr, Kindra, & Dash, 2008) of websites. Aesthetics has also been found to be one of the biggest factors in how preferable websites are perceived (Schenkman & Jönsson, 2000). Overall, however, the present study did not reveal any statistically reliable differences between either the ratings or the effectiveness of the website by website color. Nonetheless, the present study did reveal a statistically reliable interaction between color, gender and ratings. This interaction indicates that females and males differed in how the red website was rated. Inspecting the graphs of the ratings made by the females and males, a trend can be discerned where females rate the red website higher than the males on every rating. Post hoc analyses confirmed this trend with the exception of satisfaction. This partly supports the first hypothesis, that color effects the ratings of websites depending on gender.

The color red and its effect on males. The gender-dependent effects of red on males has been investigated in the literature, ranging from red sportswear being linked to higher success in combat sports (Rowe, Harris, & Roberts, 2005), to the color of red impairing performance

in males in general knowledge tests (Gnambs, Appel, & Batinic, 2010). In a somewhat similar setting to the present study, Gnambs, Appel, and Batinic (2010) investigated various colors' effects on females and males by manipulating the color of different visual parts of a general knowledge test distributed online. They found that the manipulation of the color of elements present in the general knowledge test, like progress bars and buttons to proceed, had a gender-dependent impact on performance (Gnambs, Appel, & Batinic, 2010). Males' performance was significantly impaired when the different elements appeared in red, as opposed to females who had no performance impairment at all. Likewise, Elliot et al. (2007) found that red impaired the performance of both females and males. To explain these effects on males' performance, prior studies draw on an evolutionary perspective that in non-humans, the color of red is often a sign of dominance (Gnambs, Appel, & Batinic, 2010; Ioan et al., 2007). On this basis males might feel threatened by the color of red, and so rated the red website, in the present study, lower than females.

The research presented above found differences in actual performance when exposed to red, and not on ratings like in the present study. This could possibly be explained through the experimental task used when investigating performance in this study. In the experiment by Gnambs, Appel, and Batinic (2010), participants had to answer a big amount of questions ranging over several pages. In the present study, the cognitively demanding task was finished rather quickly, and consisted of reading and comprehending text. As a general knowledge test involves answering questions drawing from your own knowledge, where the task used in the present experiment required actively reading and comprehending information, there might be different mechanisms at work between the present and the study by Gnambs, Appel, and Batinitc (2010).

The color red and its effect on females. Females rated the red website more highly than males. This indicates that females found the red website more efficient, more valuable, and had a more positive attitude towards the website than males. However, as color alone did not show any reliable effects on the website ratings, there is not much that can be said about the actual effects of red. If anything, red should be avoided in the design of website intended for use by both males and females. However, if the website is intended to be used mostly by females, replacing the color red, if it is used, with something with more gender neutral might not be as urgent.

Limitations of the Experiment

It was not possible to translate every element present in the websites used for inspiration in the present study to an experimental setting. This included elements such as references, valid news reports, contact details and pictures. The absence of these elements might have affected the realism of the experiment. However, references could not be used as that would have been incoherent with the text about the fictitious disease, which had none. Contact details for the experimenters were provided, but no contact details for the fictive authors of the experimental website was possible to provide.

Length of the experiment. As all of the scales were developed to assess participants' general impression of the website, there might be issues with how the websites were rated. To obtain more valid ratings of the complete website, extending the time of participants spent using the website may have been preferable. This to ensure that participants had navigated the website to an extent that would let them reliably make a judgment of the website as a whole. In the present study, participants were instructed to use the website to quickly search for the correct answer to the experimental question. The main purposed was to provide an as natural context as possible for interaction with the website, and so, only one experimental question was chosen. This was based on statistics from Kuehn (2013), stating that eight out of 10 visits to websites providing medical information online starts from a search engine like Google (www.google.com). This would make starting anywhere but where a relevant link from a search engine would lead you not likely to occur in reality. It was also based on the suggestions by Bonnardel et al. (2011) about investigating color in a natural context.

Due to the nature of search engines, it is not possible to effectively search for two unrelated medical issues at once, which the experimental design of prior studies assessing peoples' search behavior online attest to (Diaz, Griffith, Reinert, Friedmann & Moulton, 2002). This also contributed to the conclusion that to mimic the most common search practice for finding medical information online, it is reasonable to use only one task, even if that limited the time participants spent using the website. Moreover, as some opinions, like aesthetics, of websites have been shown to be formed and stabilize as quick as 50ms (Tractinsky, Cokhavi, Kirschenbaum & Sharfi, 2006) to 500ms (Tractinsky, Cokhavi & Kirschenbaum, 2004), there is little reason to think that the time spent using the experimental website was insufficient for the present purpose of rating the experience of using the website.

The amount of color used. Prior studies that have found statistically reliable effects of color have typically used a rather unnatural color manipulation of websites and interfaces (cf. Bonnardel et al., 2011). For example, the manipulation would affect almost every element on the website, even elements that normally would not be colored on regular websites. This includes elements such as text, and the background color of the text. This unnatural manipulation could create the feeling of a different color tone of the entire screen, rather than the feeling of a regular website but with a different color scheme, like a realistic manipulation would do. The approach of doing a full color manipulation like described above has some issues. There is research on reading comprehension using computer screens that indicate that legibility is sensitive to both the brightness and luminosity contrast between background and text-color (Gradisar, Humar, & Turk, 2007). If most elements in the website is manipulated, including the text and the background of the text, that could alter the brightness and luminosity contrast of the website material. This could lead to obtained experimental effects being due to decreased legibility of certain color combinations and contrasts, rather than actual cognitively beneficial effects elicited by certain colors.

The issues raised above, in combination with the design of the websites used for inspiration when designing the experimental website, led to the amount of the experimental website being manipulated with color. Notably, several of the participants in the initial pilot study reported that they failed to remember the color of the experimental website they used. This even though they were aware of the purpose and manipulation of the experiment. However, other studies on color and cognitive performance has found cognitively enhancing effects to be largely unconscious (Mehta & Zhu, 2009). The results of this study gives reason to questions the applicability of the potential effects of color in the context of searching for medical information online.

Future Research

Even though this study failed to show statistically significant main effects of color in the present context of presenting medical information online, there may be other contexts where color might be more apt for producing beneficial effects. The context of this study was narrow and specific, with one single task in mind. In light of earlier research about the gender-dependent effect of red on males and the methodology used in those studies (Gnambs, Appel, & Batinic, 2010; Ioan et al., 2007), it may be of interest to investigate color in a context with

another type of experimental task. As discussed, there might be different mechanisms at work comparing knowledge-based tasks, such as taking an online knowledge test, with tasks based around comprehending available information. Therefore, it might be of interest to investigate color in a context like internet based education, where knowledge and learning is a major part of the cognitive processes involved.

Color and culture in specific contexts. The literature investigating color differences culturally display rather widespread color preferences across cultures (Ou, Ronnier Luo, Sun, Hu, Chen, Guan, & Richter, 2012; Cyr, Head, & Larios, 2010; Marcus & Gould, 2000). Therefore, it might be of interest to investigate how color affects people from different cultures in a context like the one in this study. Some authors claim there is sex differences in color preferences across cultures (Hurlbert & Ling, 2007), while some claim there are no such differences (Taylor, Clifford, & Franklin, 2012). As the color of red had a disadvantageous effect on males as compared to females in this study, it might be of interest to investigate these potential differences culturally.

Perception of disease-severity. Prior research indicates that colors are able to elicit emotions and associations, such as red commonly being associated with failure, dominance or warnings (Mehta & Zhu, 2009) and blue being associated with relaxation (Gorn et al., 2004) and generally viewed as pleasing (Valdez & Mehrabian, 1994). Therefore, it would be interesting to investigate the effect of color on the perception of disease-severity. Newly diagnosed patients are perhaps influenced by the color presented along with information about their condition when researching the condition online. Since red is often associated with failure and warnings (Elliot et al., 2007; Meta & Zhu, 2009), patients might be more alarmed when learning about their condition in conjunction with red, compared to if the information would have been presented in conjunction with blue. This would be relevant information for Health care professionals in helping patients deal with new diagnoses in a constructive and calming way.

Conclusions

To sum up, this study investigated the effect of color on users' ratings and performance of using a website, by systematically manipulating the color of the website. The analysis showed a statistically significant three-way interaction of website color, gender and ratings, where the

red website was rated higher by females than males on the dimensions efficiency, value, and attitude. The negative effect of red on males as compared to females finds support in the literature, where males have been shown to perform worse than females when exposed to the color red (Gnambs, Appel & Batinic, 2010; Ioan et al., 2007). However, the literature is not particularly coherent on this issue, as red has also been shown to impair performance independent of gender (Elliot et al., 2007). Notably, however, no color dependent impairment in the website ratings of males and females was found in the present study, rather the ratings of the red website was lower by males than by females. The tentative indication is, therefore, that females value red websites more than the males.

No other statistically significant results were found. This questions the potential positive effect on cognitive functions of colors found in other studies (Bonnardel et al., 2011), at least in the context of searching for medical information online. For the future, it is of interest to conduct more research on colors in websites or other interfaces using different experimental tasks with different cognitive functions at work. Possibly the color red only impairs performance on certain types of tasks. It would also be of relevance to investigate if there is any positive effect on performance in relation to how much of the area of the screen is manipulated with color. Continued research promises to provide valuable clues for optimal website design.

References

Ajzen, I., 1989. Attitude structure and behavior. In: Pratkanis, A.R., Breckler, S.J., Greenwald, A.G. (Eds.), *Attitude Structure and Function*. Lawrence Erlbaum Associates, Inc, Hillsdale NJ, pp. 241–274.

Babin, B. J., Hardesty, D. M., & Suter, T. A. (2003). Color and shopping intentions: the intervening effect of price fairness and perceived affect. *Journal of Business Research*, *56*(7), 541-551.

Bellizzi, J. A., & Hite, R. E. (1992). Environmental color, consumer feelings, and purchase likelihood. *Psychology & Marketing*, *9*(5), 347-363.

Bonnardel, N., Piolat, A., & Le Bigot, L. (2011). The impact of colour on Website appeal and users' cognitive processes. *Displays*, *32*(2), 69-80.

Cohn, R. (1894). Ueber einen in den thierischen Geweben sich vollziehenden Reductionsprocess. *Zeitschrift für physiologische Chemie*, *18*(2), 133-136.

Crowley, A. E. (1993). The two-dimensional impact of color on shopping. *Marketing letters*, *4*(1), 59-69.

Cyr, D., Head, M., & Larios, H. (2010). Colour appeal in website design within and across cultures: A multi-method evaluation. *International journal of human-computer studies*, *68*(1), 1-21.

Cyr, D., Kindra, G. S., & Dash, S. (2008). Web site design, trust, satisfaction and e-loyalty: the Indian experience. *Online Information Review*, *32*(6), 773-790.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: a comparison of two theoretical models. *Management science*, *35*(8), 982-1003.

Diaz, J. A., Griffith, R. A., Ng, J. J., Reinert, S. E., Friedmann, P. D., & Moulton, A. W. (2002). Patients' use of the Internet for medical information. *Journal of general internal medicine*, *17*(3), 180-185.

Elliot, A. J., & Aarts, H. (2011). Perception of the color red enhances the force and velocity of motor output. *Emotion-APA 11*(2), 445.

Elliot, A. J., Maier, M. A., Moller, A. C., Friedman, R., & Meinhardt, J. (2007). Color and psychological functioning: The effect of red on performance attainment. *Journal of experimental psychology. General*, *136*(1), 154.

Eysenbach, G., & Köhler, C. (2002). How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. *BMJ: British Medical Journal*, *324*(7337), 573.

Flavián, C., Guinalíu, M., & Gurrea, R. (2006). The role played by perceived usability, satisfaction and consumer trust on website loyalty. *Information & Management*, *43*(1), 1-14.

Gilmour, J. A., Huntington, A., Broadbent, R., Strong, A., & Hawkins, M. (2012). Nurses' use of online health information in medical wards. *Journal of Advanced Nursing*, *68*(6), 1349-1358.

Gnambs, T., Appel, M., & Batinic, B. (2010). Color red in web-based knowledge testing. *Computers in Human Behavior*, *26*(6), 1625-1631.

Gorn, G. J., Chattopadhyay, A., Sengupta, J., & Tripathi, S. (2004). Waiting for the web: how screen color affects time perception. *Journal of marketing research, 41*(2), 215-225.

Gradisar, M., Humar, I., & Turk, T. (2007, June). The legibility of colored web page texts. In *Information Technology Interfaces, 2007. ITI 2007. 29th International Conference on* (pp. 233-238). IEEE.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*(2), 95-112.

Hilton, S., & Smith, E. (2010). Public views of the UK media and government reaction to the 2009 swine flu pandemic. *BMC Public Health*, *10*(1), 697.

Hurlbert, A. C., & Ling, Y. (2007). Biological components of sex differences in color preference. *Current Biology*, *17*(16), R623-R625.

Ioan, S., Sandulache, M., Avramescu, S., Ilie, A., Neacsu, A., Zagrean, L., & Moldovan, M. (2007). Red is a distractor for men in competition. *Evolution and Human Behavior*, *28*(4), 285-293.

Krewski, D., Lemyre, L., Turner, M. C., Lee, J. E., Dallaire, C., Bouchard, L., ... & Mercier, P. (2006). Public perception of population health risks in Canada: Health hazards and sources of information. *Human and ecological risk assessment*, *12*(4), 626-644.

Kuehn, B. M. (2013). More Than One-Third of US Individuals Use the Internet to Self-diagnose. *JAMA*, *309*(8), 756-757.

Lichtenfeld, S., Maier, M. A., Elliot, A. J., & Pekrun, R. (2009). The semantic red effect: Processing the word red undermines intellectual performance. *Journal of Experimental Social Psychology*, *45*(6), 1273-1276.

Lichtle, M. (2007). The effect of an advertisement's colour on emotions evoked by an ad and attitude towards the ad. *International Journal of Advertising*, *26*(1), 37.

Maga, J. A. (1973). Influence of freshness and color on potato chip sensory preferences. *Journal of Food Science*, *38*(7), 1251-1252.

Marcus, A., & Gould, E. W. (2000). Crosscurrents: cultural dimensions and global Web user-interface design. *Interactions*, *7*(4), 32-46.

Mehta, R., & Zhu, R. J. (2009). Blue or red? Exploring the effect of color on cognitive task performances. *Science*, *323*(5918), 1226-1229.

Norman, D.A. (2002[1988]). *The design of everyday things*. (First Basic paperback ed.) New York: Basic Books.

O'Connor, J. B., & Johanson, J. F. (2000). Use of the Web for medical information by a gastroenterology clinic population. *JAMA: the journal of the American Medical Association*, *284*(15), 1962-1964.

O'Connor, M., Whitfield, T. W. A., & Wiltshire, T. J. (1989). Architects' colour terminology: Is there order?. *Color Research & Application*, *14*(2), 79-85.

Ou, L. C., Ronnier Luo, M., Sun, P. L. et al. (2012). A cross-cultural comparison of colour emotion for two-colour combinations. *Color Research & Application*, *37*(1), 23-43.

Rice, R. E. (2006). Influences, usage, and outcomes of Internet health information searching: multivariate results from the Pew surveys. *International journal of medical informatics*, *75*(1), 8-28.

Robins, D., & Holmes, J. (2008). Aesthetics and credibility in web site design. *Information Processing & Management*, *44*(1), 386-399.

Rowe, C., Harris, J. M., & Roberts, S. C. (2005). Sporting contests: seeing red? Putting sportswear in context. *Nature*, *437*(7063), E10.

Schackel, B., 1991. Usability—context, framework, design, and evaluation. In: Shackel, B., Richardson, S. (Eds.), *Human Factors for Informatics Usability*. Cambridge University Press, UK, pp. 21–38.

Schenkman, B. N., & Jönsson, F. U. (2000). Aesthetics and preferences of web pages. *Behaviour & Information Technology*, *19*(5), 367-377.

Schmidt, K. E., Liu, Y., & Sridharan, S. (2009). Webpage aesthetics, performance and usability: Design variables and their effects. *Ergonomics*, *52*(6), 631-643.

Suziedelyte, A. (2012). How does searching for health information on the Internet affect individuals' demand for health care services?. *Social Science & Medicine, 75*(10), 1828–1835.

Taylor, C., Clifford, A., & Franklin, A. (2012). Color Preferences Are Not Universal. *Journal of Experimental Psychology: General*, Nov 12, 2012.

Teo, H. H., Oh, L. B., Liu, C., & Wei, K. K. (2003). An empirical study of the effects of interactivity on web user attitude. *International Journal of Human-Computer Studies*, *58*(3), 281-305.

Tractinsky, N., Cokhavi, A., & Kirschenbaum, M. (2004, December). Using ratings and response latencies to evaluate the consistency of immediate aesthetic perceptions of web pages. In *Third Annual Workshop on HCI Research in MIS, Washington, DC* (pp. 40-44).

Tractinsky, N., Cokhavi, A., Kirschenbaum, M., & Sharfi, T. (2006). Evaluating the consistency of immediate aesthetic perceptions of web pages. *International Journal of Human-Computer Studies*, *64*(11), 1071-1083.

Valdez, P., & Mehrabian, A. (1994). Effects of color on emotions. *Journal of Experimental Psychology: General*, *123*(4), 394.

Whitfield, T. W., & Wiltshire, T. J. (1990). Color psychology: A critical review. *Genetic, social, and general psychology monographs*, *116*(4), 385.