Colour appeal in website design within and across cultures: A multi-method evaluation

Dianne Cyr, Milena Head, Hector Larios

Faculty of Business, Simon Fraser University, 15th Floor, Central City Tower, 13450 102nd Avenue, Surrey, BC, Canada V3T 5X3
DeGroote School of Business, McMaster University, Hamilton, Ontario, Canada
School of Interactive Arts and Technology, Simon Fraser University, Surrey, BC, Canada

Received 29 July 2009; accepted 21 August 2009
Communicated by S. Wiedenbeck
Available online 28 August 2009

Abstract

Colour has the potential to elicit emotions or behaviors, yet there is little research in which colour treatments in website design are systematically tested. Little is known about how colour affects trust or satisfaction on the part of the viewer. Although the Internet is increasingly global, few systematic studies have been undertaken in which the impact of colour on culturally diverse viewers is investigated in website design. In this research three website colour treatments are tested across three culturally distinct viewer groups for their impact on user trust, satisfaction, and e-loyalty. To gather data, a rich multi-method approach is used including eye-tracking, a survey, and interviews. Results reveal that website colour appeal is a significant determinant for website trust and satisfaction with differences noted across cultures. The findings have practical value for web marketers and interface designers concerning effective colour use in website development.

Keywords: Colour appeal; Culture; Website design; Multi-methodology; Eye-tracking

1. Introduction

Colour is everywhere around us. Our visual and cognitive systems have adapted to perceive and process colour information, which is contained in every visual stimulus we encounter (Elliot and Maier, 2007; Kaya and Epps, 2004). Colour information such as hue, brightness, and saturation has the potential to affect our perceptions, physiological reactions, emotional reactions or behavioral intentions (Valdez and Mehrabian, 1994). In the commercial realm, colour influences our attitude and expectations toward brands. For example, red is the colour symbolizing Coca Cola, and blue is associated with IBM. While there has been some investigation of colour in print or other media, “despite its importance surprising little is known about the influence of colour in advertising” (Latomia and Happ, 1987, p. 37). Further, there is relatively little research regarding colour in online, Internet-based environments. In particular, the impact of colour in website design is sparse although a few researchers (Lui et al., 2004) have tackled this challenge.

Although research into the psychology of colour is not a well developed area, several studies have provided evidence of a relationship between colours and emotions. Long-wavelength colours (e.g., red and yellow) have been hypothesized to be more negatively arousing than short-wavelength colours (e.g., blue and green) (Jacobs and Hustmyer, 1974; Wilson, 1966), and these physiological reactions seem to be indicative of various psychological outcomes such as anxiety or pleasure (Kwallek et al., 1988;
Valdez and Mehrabian, 1994). Emotional reactions to colours may be influenced by physiological and environmental variables that have not yet been analyzed in most studies. On the one hand, some argue that our emotional reactions to colour have an evolutionary and therefore biological origin (Elliot and Maier, 2007). Wells et al. (2008), for example, showed that gorillas and chimpanzees attended more to blue and green-coloured stimuli than to red. While one should be cautious to generalize to human populations, it is worth noting that these results are consistent with findings in human studies. On the other hand, variables like age, sex, and culture are also thought to influence emotional and behavioral reactions to colour. As Valdez and Mehrabian (1994) pointed out, research on the effect of colour on emotions has been inconclusive and further research is needed to clarify the issues.

Simon (2001) suggested that website design which may include elements such as colour, shapes, images, or streaming video contributes to a user’s perception of the website. He further adds that such perception has theoretical underpinnings in communication theory (Pan et al., 2004) and relates to an effective website interface. In this vein, and as part of our theoretical underpinning for the current research, we more specifically refer to a cognitive-affective model of organizational communication for designing information technology as elaborated by Te’eni (2001). The communication medium (which is a website in our case) has the potential to affect an impact on the user (including various affective reactions such as trust). Consumer oriented websites that match the social and emotional perceptions of users are expected to increase trust and be more engaging (Pieters et al., 2002). Further, recent HCI research has emphasized the importance of hedonic or emotional elements in website design as important to user enjoyment or loyalty (Cyr et al., 2009).

Although colour has the potential to elicit emotions or behaviors, there are few studies in which various colour treatments in website design are tested regarding their impact on the viewer’s outcome state such as trust or satisfaction. However, Cyr (2008) found that visual design of the website (which includes colour) resulted in trust, satisfaction, and loyalty. Further, Kim and Moon (1998) examined colour for a cyber banking interface and discovered that colour had a main effect on the trustworthiness of the interface. Although there is little theory that has previously been used to explain how colour influences user reactions on the Web, in this study we draw on Walters et al. (1982) theory of psychological reversals to colour to help explain user reactions to colour in an online context. Further, this theory of colour is married with the cognitive-affective model of communication as mentioned above to gain a fuller appreciation of how colour elicits emotion in the user.

To our knowledge there is no study in which colour is considered as a catalyst to user loyalty toward a website. In this sense, if a user likes the colour of a website, which arouses reactions of trust or satisfaction, then one would expect the user to return to the website or to purchase from it in the future, typically termed e-loyalty (Falk and Miller, 1992). Evocation of this potential flow of relationships has practical value if web marketers and interface designers wish to use colour as a catalyst in website development with a goal of attracting potential online consumers both in local and international markets. Especially rare are rigorous studies of colour in website design across cultures. This area represents an important and overlooked topic. According to Noiwan and Norcio (2006):

**Empirical investigations on the impacts of cultural factors on interface design are absolutely vital... Interface designers need to understand colour appreciation and colour responses of people in different cultures and regions” (pp. 103, 104).**

Tractinsky (2004) further emphasizes that theory building and systematic testing are critical to understand cross-cultural issues, but to date much of the work in this area has been based on personal experiences or case studies (Marcus and Gould, 2000).

To address these apparent gaps in research to date, colour appeal in website design is the focus of the current investigation. Colour appeal is defined as the degree to which colours on websites are perceived by the user as pleasing, appealing, and appropriate. Related to colour appeal, in this research, the following questions are examined:

1. How do different applications of colour (blue, yellow, grey) in website design influence colour appeal of the website? And in turn, does increased colour appeal impact important e-loyalty antecedents such as trust and satisfaction?
2. Between diverse cultural groups (Canada, Germany, Japan) are there differences in how applications of colour on websites impact colour appeal? Does this in turn impact user trust and satisfaction?

To gain a fuller understanding of these complex phenomena, a multi-method approach is undertaken in which data is collected using an eye-tracking device, along with surveys and interviews. This methodology is in alignment with Pieters et al. (2002), and more recently Noiwan and Norcio (2006) who examined cultural differences in the perception of animated graphics and recommend the use of eye-tracking for detailed research to understand how users interact with specific interface elements. Also from a methodological perspective, a new construct for colour appeal is created and successfully validated. The scale focuses specifically on colour in website design rather than examining general design characteristics of the website. Rigor in methodology and the development of new metrics in e-commerce align well with the goal of measuring e-commerce elements in network-enabled organizations (Short et al., 1976). This especially
applies to theoretical development of specifics of visual design, including the metric of colour appeal as it impacts the international online consumer.

2. Theoretical framework and research model

A cognitive-affective model of communication for designing information technology (Te’eni, 2001) posits that the communication process, which includes the communication medium and the message form, will have an impact on how the communication is received. Characteristics of the physical medium on which the message is transmitted influence the “relationship” which in the context of the model includes whether or not “the communication act is judged to be trustworthy and appropriate” (p. 257). With reference to the present research, the website is the communication medium, and the message form is characterized by differences in colour application with potential to influence the relationship an online user has with the website. The relationship may be one that is trusting and satisfied if the message is viewed as appealing, or alternately neither of these affective reactions may occur if the message fails to establish a positive relationship with the user.

There is a possibility for cultural contextualization of this theory. In line with Habermas (1984), communication action is situated in the context of social norms and values. As part of the communication process individuals or groups have expectations regarding the representation of the message. For example, some cultures expect to read a message right to left, while others from left to right. The same could be said for expectations of colour (Barber and Badre, 1998). Representations incompatible with user expectations require additional effort and are therefore rejected in favor of more compatible representations (Barber, 1988).

The idea that colour may affect psychological processes dates back to early work by Goldstein (1939). Supporting the view that colour has the ability to impact user psychological reactions Walters et al. (1982) proposed a theory of psychological reversals. Within the scope of the theory it is asserted there are two levels of preferred felt arousal – one high and one low. An individual may be in one of the two stable psychological mood states at a time, yet can reverse to the other state. Some colours serve to arouse and excite an individual, while other colours elicit relaxation. This assumption was tested in a work environment in which participants were asked to make colour preference choices. The results of the study strongly supported the association between colour preference and arousal. This theoretical work gains support in an online context with reference to research by Kim and Moon (1998) on cyber banking. With a focus on design elements such as clip art and colour, the colour layout of the interface was found to be important to customer perceptions of trustworthiness. More specifically, cool tones such as blue were preferred over warm tones such as red or yellow. Relevant to the current investigation, we would expect different colours to have different levels of appeal to users, hence resulting in differing levels of trust or satisfaction.

In sum, we propose that the cognitive-affective model of communication emphasizes the medium of communication through which a user perceives the website. This in turn results in attributions of the website and whether or not is trustworthy. This theory is complementary to the theory of psychological reversals in which colour has been shown to play a major role in the psychological arousal of the user. Related to these theories we would expect colour treatments to appeal to the user to varying degrees, which in turn results in trust or satisfaction based on user perceptions as to whether a colour treatment is deemed appropriate. This would further result in whether users exhibit loyal behavior toward a website.

Based on these theories and assumptions we offer our research model. The model is developed to test the impact of website colour appeal on online trust and satisfaction which ultimately result in online loyalty, or e-loyalty. Trust and satisfaction were chosen as key antecedents to loyalty based on the establishment of these relationships in other work related to website design (Cyr, 2008; Cyr et al., 2008, 2007; Flavián et al., 2006; Lam et al., 2004). Further the model explores the influence of various website colour schemes (grey, blue, yellow) on colour appeal, trust and satisfaction. The proposed research model appears in Fig. 1. In addition to testing the model relationships for combined national cultural groups, between-group cultural differences are also explored. Specifically, we seek to understand if culture moderates the relationships between colour appeal and e-loyalty antecedents (trust and satisfaction), and if cultural groups vary in their reactions to various website colour schemes.

A review of the literature that frames the research constructs and hypotheses for testing is presented in the following sections. This is followed by the methodology used in the investigation, analysis and results, and a

![Fig. 1. Proposed research model.](image-url)
discussion of both theoretical and practical implications of this research.

3. Hypothesis development

3.1. Colour appeal, trust and satisfaction

The study of colour and colour appeal has interdisciplinary connections. Colour has been studied by a variety of researchers, from artists to zoologists, with topics focused on colour in art over the ages (Fornell, 1982), visual perception and explanations of colour categories (Gorn et al., 1997) or naming of colours (Bellizzi and Hite, 1992). The HCI literature has focused on colour and symbolism (Bagozzi and Yi, 1989), although the work is not empirically tested. It is interesting that there are few, if any, theoretical models that expand knowledge on colour across cultures – particularly in the context of website design.

For many years psychologists have been interested in the effect of colour on preferences (Goldberg et al., 2002). “Colours are known to possess emotional and psychological properties” (Lichtle, 2007, p. 91) and have the potential to convey commercial meaning in products, services, packaging, and Internet design. Marketers have long known the power of colour in corporate brand-building for logos or displays (Lui et al., 2004; Rivard and Huff, 1988).

Cooler colours such as blue and green are generally viewed more favorably than warmer colours such as yellow or red (Goldberg and Kotval, 1999; Latomia and Happ, 1987; Marcus and Gould, 2000). More specifically, blue is generally associated with “wealth, trust, and security” (Lichtle, 2007) and is universally liked (Carte and Russell, 2003; Meyers-Levy and Peracchio, 1995; Nielsen and Del Galdo, 1996). In part, this may explain the use of blue by corporate entities such as banks (at least in North America) to establish a professional and credible image. Alternatively, orange denotes “cheapness” (Lichtle, 2007). Other research on colour suggests hue (as in primary colours red, blue, yellow), brightness (light colours such as white versus dark colours such as black or grey), and saturation (intense versions of a colour versus pastels) all have an effect on individual perceptions (Latomia and Happ, 1987). Colour likewise has an influence on behavioral intention, with blue producing stronger buying intention than red (Becker, 2002; Latomia and Happ, 1987). These findings, mostly from marketing and advertising, are considered here in the context of website design.

In an online environment, elements of website design include colour, images, shapes, use of photographs among other characteristics and these are anecdotally expected to provide the user with emotional appeal, a sense of the aesthetic, or a positive impression of the overall graphical look of a website (Cyr et al., 2005; Everard and Galletta, 2006; Fornell and Larcker, 1981; Rousseau et al., 1998). Although there is no shortage of empirically based studies on website design characteristics, these studies consider general characteristics of visual design of the website (Cyr and Trevor-Smith, 2004) or general usability (Falk and Miller, 1992; Madden et al., 2000; West and Hepworth, 1991) and do not usually address the specific impact of colour on user perceptions such as trust or satisfaction in website design.

With respect to trust, researchers have endeavored to uncover the complexities of trust in online environments (Berlin and Kay, 1969; Chattopadhayay et al., 1999; Cyr, 2008; Del Galdo and Neilson, 1996; Holland and Baker, 2001; Loftus and Mackworth, 1978; Nunally, 1978, Venkatesh et al., 2003).¹ Corritore et al. (2003) provide a definition of online trust that includes cognitive and emotional elements, with trust encompassing “an attitude of confident expectation in an online situation or risk that one’s vulnerabilities will not be exploited” (p. 740). Online trust relates to consumer confidence in a website and willingness to rely on the vendor in conditions where the consumer may be vulnerable to the seller (Holland and Baker, 2001). Similar to the preceding definition of trust, in this investigation online trust refers to general trust on the part of the user for the website, including information presented and the transaction process.

Website usability can significantly impact trust (Flavián et al., 2006). In a mixed country sample, Cyr (2008) found that visual design leading to trust was significant (p < .001). Since colour is one element of visual design, we would argue it is reasonable to test if the more specific construct of colour appeal has a relationship to trust. Further, research into aesthetics or beauty of the website found high correlations between perceived aesthetics and perceived ease of use (Lavie and Tractinsky, 2004; Tractinsky, 1997, 2000; van der Heijden, 2003). In turn, for web users perceived ease of use is known to influence trust (Gefen et al., 2003). Karvonen (2000) suggests that a direct relationship exists between trust and the “aesthetic beauty” of a website. Adapted from marketing and as noted above, certain colours such as blue are perceived as more aesthetically appealing and are associated with trust (Lichtle, 2007), while other colours are perceived by the user as less favorable. As Simon (2001) suggests, effectiveness of the communication medium through website design including colour can result in trust and more positive perceptions of the website. In one study in which colour and online trust are explicitly considered when banking online, a preferred colour layout of the customer interface was found to significantly impact whether or not the website was considered trustworthy (Kim and Moon, 1998).

¹A thorough review of trust in offline and online settings is not feasible within the scope of the present paper. However, the reader may wish to refer to Rousseau et al. (1998) for a comprehensive and cross-disciplinary critique of offline trust. Online trust is comprehensively examined by Gefen et al. (2003) who provide a summary of previous conceptualizations of trust, as well as an integrated model of trust in online shopping. In addition, Corritore et al. (2003) offer a useful overview of online trust.
The preceding leads to our first hypothesis.

**Hypothesis 1.** For all cultural groups, increased colour appeal will result in greater trust.

Similar to trust, if a website user finds colours on the website visually appealing then this is likely to lead to satisfaction. Satisfaction on the Web relates to “stickiness” and “the sum of all the website qualities that induce visitors to remain at the website rather than move to another site” (Hoffman and Novak, 1996). Artifacts of website design that contribute to satisfaction are numerous and varied. Palmer (2002) validated design metrics for websites and found site organization, information content and navigation important to website success, including intent to return to the site. In other research, website design and the “ambience associated with the site itself and how it functions” is an antecedent to satisfaction (Straub, 1989).

In alignment with other researchers (Agarwal and Venkatesh, 2002; Cronbach, 1971; Falk and Miller, 1992) in this investigation website satisfaction refers to overall contentment with the online experience.

Numerous researchers have determined that an effectively designed website including the use of appropriate colours may engage and attract consumers resulting in satisfaction with an online vendor (Adams and Osgood, 1973; Fernandes, 1995; Hair et al., 1995; Karvonen, 2000; Moore et al., 2005; Straub, 1989). In a mixed country sample, Cyr (2008) found that visual design leading to satisfaction was significant \( p < .001 \). As with trust, since colour is one element of visual design we are interested to see if colour appeal is related to satisfaction in the current investigation. In one study, ineffective colour combinations on graphic design impeded user performance and satisfaction (Koufaris, 2002). Thus, we propose:

**Hypothesis 2.** For all cultural groups, increased colour appeal will result in greater satisfaction.

3.2. Trust, satisfaction, and e-loyalty

In online settings, “understanding how or why a sense of loyalty develops in customers remains one of the crucial management issues of our day” (Lam et al., 2004, p. 156). Online shoppers are more likely to revisit a website if they like its design and capabilities (Falk and Miller, 1992; Junglas and Watson, 2004; Madden et al., 2000; Venkatesh and Ramesh, 2006). Despite the apparent importance of developing loyal relationships with customers, “there is limited academic research on the relative importance on individual elements of website design and their effect on customer loyalty” (Madden et al., 2000, p. 99). Based on previous research, online loyalty or e-loyalty has been conceived as a “consumer’s intention to buy” from a website, and that consumers will not change to another website (Falk and Miller, 1992). Schjins (2003) refers to attitudinal loyalty as commitment of a customer to a vendor including “likelihood of future usage”. In this study, e-loyalty is defined as future intention on the part of the user to visit the website again or to purchase from it in the future.

Yoon (2002) tested the relationship of website properties and navigation functionality to website trust, which in turn was predicted to result in on/offline purchase intention (similar to e-loyalty). Results indicated website properties and trust are related and influence online purchase intentions. In other works, perceived website usability directly influenced online consumer trust, which in turn influenced e-loyalty (Cronbach, 1971; Falk and Miller, 1992). While the relationship of trust to loyalty has been previously confirmed as noted here, in the current study it is of interest to consider this relationship in the specific context of differently coloured websites.

**Hypothesis 3.** For all cultural groups, higher trust will result in greater e-loyalty.

Repeated satisfaction with a vendor eventually results in e-loyalty (Cronbach, 1971; Falk and Miller, 1992; Kim and Benbasat, 2006; Lankford, 2000). As with trust, although the relationship of satisfaction to loyalty has been previously confirmed, it is now tested in the context of three colour treatments for the websites.

**Hypothesis 4.** For all cultural groups, increased satisfaction will result in greater e-loyalty.

3.3. Culture, colour appeal, trust, and satisfaction

In the literature on culture, colour, and cognition there exist two opposing views. A Universalistic view prescribes pan-human cognitive processing of colour naming and colour perception. Alternately, Cultural Relativism suggests that colour perception is mostly shaped by culturally specific language associations and perceptual learning (Berlin and Kay, 1969; Kay et al., 1991). In the realm of the Internet, the prevailing view is more aligned to Cultural Relativism with an emphasis on website localization (Barber and Badre, 1998; Cyr and Trevor-Smith, 2004). That is, users are expected to prefer websites that are adapted to their own culture. When localizing a website, in addition to language translation, details such as currency, colour sensitivities, product or service names, images, gender roles, and geographic examples are considered.

Culture has implications for Internet usage and affects e-commerce trust (Cronbach, 1971; Garrett, 2003; Holland and Baker, 2001), Internet marketing (Straub et al., 2004), and website development (Hu et al., 2004; Singh et al., 2003). Web interface acceptance and preferences for design features differ between cultures (Bagozzi and Yi, 1989; Boor and Russo, 1993; Cyr et al., 2009). Depending on a user’s culture there are different preferences for design elements including colour (McKnight et al., 2002; Lin et al., 2005; Singh et al., 2003).

As noted by Gefen (2000), there is a need to conduct research on cross-cultural effects and trust. In studies in
which trust and culture are considered, the results are often mixed or inconclusive (Hofstede, 1980; Holland and Baker, 2001; Lavie and Tractinsky, 2004). In an investigation in which culture and trust were examined for Canadian, American, German, and Japanese participants, similar perceptions of a local website were found concerning trust for Canadians and Americans, while there was a modest difference \( p < .1 \) for Americans with Germans, and significant differences \( p < .01 \) for Americans, Canadians and Germans with Japanese (Cyr et al., 2005).

Further, Cyr (2008) modeled visual design of the website to trust, satisfaction, and loyalty across cultures. Partial Least Square (PLS) analyses results indicated that with the overall sample, visual design leading to trust was significant \( (p < .001) \), visual design leading to satisfaction was significant \( (p < .001) \) and both trust and satisfaction leading to e-loyalty were significant \( (p < .001) \). In subsequent PLS analyses, there were differences as to whether users trusted and were satisfied with visual design moderated by culture. As mentioned earlier, since colour is one element of visual design, we would argue that it is reasonable to test if the more specific construct of colour appeal has relationships to trust and satisfaction moderated by culture.

As an element of communication, colour can elicit affect in the user. In a study focused on determining the meaning and preferences for 10 colours across 8 countries (Madden et al., 2000), blue was generally ascribed meaning as “peaceful” or “calming”, or in the case of Americans as “pleasant”. In contrast, brown and black had associations of “sad” and “stale” across cultures. Other colours, such as yellow and orange, showed less cross-cultural consistency in terms of how they were perceived. The authors conclude, “Underlying these colour selection decisions is the notion that different combinations of colours may be capable of evoking different consumer reactions” (p. 102).

Colour connotes different meaning in different cultures (Baggozi and Yi, 1989; Barber and Badre, 1998; Rousseau et al., 1998). Red means happiness in China but danger in the United States. Colour can impact user expectations as well as overall satisfaction (Koufaris and Hampton-Sosa, 2004; Meyers-Levy and Peracchio, 1995). Building on previous work, we expect that consumer reactions to website colour schemes will result in different levels of colour appeal for different cultural groups. In turn, colour appeal as related to trust and satisfaction are expected to be moderated by culture. These assumptions are based on the cognitive-affective model of organizational communication and the theory of psychological reversal as already outlined earlier in this paper. More specifically, a cultural group will have expectations regarding the representation of the message, which include colour applications (Barber and Badre, 1998). Whether or not the applied colour is aligned with user preferences will influence psychological and affective perceptions of the website. Thus, we posit the following hypotheses:

**Hypothesis 5a.** The influence of colour appeal on trust will be moderated by culture.

**Hypothesis 5b.** The influence of colour appeal on satisfaction will be moderated by culture.

Following from the preceding, it is expected that the use of specific colours on websites can also impact colour appeal, trust and satisfaction across cultures. Cyr and Trevor-Smith (2004) evaluated the use of colour on 30 municipal websites in each of Germany, Japan, and the United States (90 websites total). Colours used on a website were matched to a colour wheel and assigned a numerical value by independent raters based on the percentage of the page on which a colour appears.\(^2\) Fifteen colours were used across the websites. Relevant to the countries in this investigation, blue was most popular on German websites, while grey was the colour most often appearing on American websites. Japanese are known to prefer brighter colours such as yellow (also supported by Noiwan and Norcio 2006). Boor and Russo (1993) noted that yellow connotes grace and nobility. In the current investigation, we extrapolate from the work by Cyr and Trevor-Smith, and Boor and Russo to examine these colour combinations (grey, blue and yellow) related to our dependent variables of colour appeal, trust, and satisfaction. Between cultural groups, different colour schemes are expected to have varying impacts on colour appeal for Japanese, German, and Canadian groups.

This leads to the final set of hypotheses:

**Hypothesis 6a.** Most appealing colours will be yellow for Japanese, blue for Germans, and grey for Canadians.

**Hypothesis 6b.** Most trusted colours will be yellow for Japanese, blue for Germans, and grey for Canadians.

**Hypothesis 6c.** Most satisfying colours will be yellow for Japanese, blue for Germans, and grey for Canadians.

### 4. Methodology

#### 4.1. Participants

Ninety participants from three countries (30 each from Canada, Germany, and Japan) took part in a laboratory experiment. These countries were chosen based on Hofstede (1980) for their known cultural diversity from each other. Refer to Table 1.

---

\(^2\)The Cool Ruler application tool was used and allows the researcher to measure the page and the sections of the page. The section of the page with a certain colour is then divided by the total page size to determine the percentage of colour on the page. The ruler appears on the screen and can measure in inches, pixels, and centimeters. Measurements are made by adjusting the ruler to either a horizontal or vertical orientation and then using the ruler as if on a sheet of paper. The Cool Ruler application is available at [http://www.softpedia.com/get/Desktop-Enhancements/Other-Desktop-Enhancements/Cool-Ruler.shtml](http://www.softpedia.com/get/Desktop-Enhancements/Other-Desktop-Enhancements/Cool-Ruler.shtml).
German and Japanese participants were recruited from international language schools located in a metropolitan area in Canada. All international participants had lived in Canada for less than one year and were over the age of seventeen. Canadian participants were recruited from a university in the same city. Table 2 summarizes the demographic profile of the study participants. The average age was consistent across cultures, there were more females in this study than males and the Canadians appeared to have more online experience. There were some outliers within the Canadian and German samples in terms of extreme number of hours spent online per week and the number of online purchases. Removing these outliers, ANOVA results showed that demographic differences did not impact the relationships being studied in this investigation.

4.2. Experimental task and design

The laboratory experiment was conducted in a controlled setting where participants browsed an e-commerce website in a usability laboratory. The study was designed as a repeated measures two-factorial experiment with three levels for each factor. The first factor was culture, where 30 participants were carefully selected to represent each of the three culture levels as characterized by nationality (Canadian, German, and Japanese). The second factor was website design, where three versions of an e-commerce website featuring electronics were created to represent three colour schemes (grey, blue, and yellow).

It should be noted that Callahan (2005) examined cultural similarities and differences in the design of university websites for eight countries, including consideration of colour. Various design predictions were made related to Hofstede’s (1980) four cultural dimensions. Callahan concluded that of 14 hypotheses only two were supported. In the current study, we therefore take a different approach, and do not aim to predict colour appeal related to Hofstede’s dimensions. Instead, we focus on colour appeal in three countries that are culturally diverse and explore possible linkages between colour schemes used in the context of culture, with our three dependent variables of colour appeal, trust, and satisfaction.

Each participant viewed each of the three colour conditions for the local country website. Thus the overall sample size for the study is 270 (30 participants by 3 countries by 3 colour treatments). Considerations for pooling repeated measures data and appropriateness of this technique in the current experiment are outlined in Appendix A. The order in which the website conditions were viewed was randomly assigned to eliminate possible order effects.

The SonyStyle website was the basis for the experimentally manipulated website and was chosen only after an extensive search by usability experts for an e-commerce website that demonstrated localization of Web content for Canada, Germany, and Japan. Localized websites were chosen for this experiment to match the expectations of participants. This was to prevent the results from being confounded by an unnatural or jarring design that is not well suited to the culture’s expectations.

Three versions of the SonyStyle website were downloaded to a local PC for experimental manipulation and to ensure consistent access speeds across subjects. The

<table>
<thead>
<tr>
<th>Country cultural dimensions.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Country dimension</strong></td>
</tr>
<tr>
<td>Power distance</td>
</tr>
<tr>
<td>Uncertainty avoidance</td>
</tr>
<tr>
<td>Masculinity</td>
</tr>
<tr>
<td>Individualism</td>
</tr>
<tr>
<td>From Hofstede (1980).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Participant demographics.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic</strong></td>
</tr>
<tr>
<td>Mean age</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Male: 7</td>
</tr>
<tr>
<td>Female: 23</td>
</tr>
<tr>
<td>Education level</td>
</tr>
<tr>
<td>High school: 10</td>
</tr>
<tr>
<td>Undergrad: 15</td>
</tr>
<tr>
<td>Grad degree: 5</td>
</tr>
<tr>
<td>Mean hours online per week</td>
</tr>
<tr>
<td>Mean # online purchases</td>
</tr>
<tr>
<td>Has purchased Sony products</td>
</tr>
<tr>
<td>No: 7</td>
</tr>
<tr>
<td>Visited tested website before</td>
</tr>
<tr>
<td>No: 19</td>
</tr>
</tbody>
</table>
SonyStyle name was removed from the websites to avoid branding effects. Each of the three country websites was designed into three versions to represent the colour manipulations (grey, blue, and yellow). These colours were chosen primarily based on research of website design (as outlined earlier) by Cyr and Trevor-Smith (2004) and Boor and Russo (1993). The primary colour manipulation zones were the left navigation bar and top graphic. These manipulation zones were approximately consistent across the localized websites with some slight variations due to differences in localized website design for the three cultures. A pilot study was conducted with five participants to ensure the three website conditions were manipulated sufficiently and appropriately. Participants in the pilot study were asked to comment on differences between the conditions, and independently and unanimously agreed there were distinct colour themes across the conditions. Sample experimental websites are shown in Appendix B.

Participants individually arrived at the usability lab and were given a brief written introduction to the research and provided a consent form to sign. For German and Japanese groups, all documents were translated and back-translated to verify accuracy, and were presented to participants in their own language. Following an introduction, each participant was fitted with the eye-tracking equipment. Task instructions provided to participants in their own language are as follows:

You will now be presented with various web pages. There is no time limit for this task. Please scan the pages with your eyes, but do not click on any links. Once you have finished scanning the page, close the browser window and inform the researcher that you have finished the task.

After participants viewed each website condition (randomly ordered), a paper-based survey was administered. At the conclusion of the session, participants were interviewed, completed a demographics form, were debriefed, and the researcher answered any remaining questions. An average experimental session lasted one hour. Participants received a $20 honorarium as compensation for their time.

4.3. A multiple method approach

Following the methodology used by Cyr et al. (2009), in this study three research and analysis methodologies are employed (eye-tracking, survey, and interview) to investigate the hypotheses. The three methods are complementary yet offer different forms of data. Survey data affords the ability to model the data to test relationships as posed in the investigation. The eye-tracking device gathers unconscious data as to exactly where a participant fixates on a website page, and for how long. Since this field is relatively new, interview data permits greater understanding as to “why” a participant responded to the survey as he or she did, and potentially the meaning behind where one actually looked on the web page as determined by the eye-tracker.

4.3.1. Eye-tracking

Much of our interaction with the world is visual in nature. We attend to and process this information using attentional and cognitive mechanisms that bring our focus of awareness on the relevant stimuli (Rayner, 1998; Recarte and Nunes, 2000). Shifts in attention can be voluntary or purposeful and they can also be involuntary or reflexive (Posner et al., 1980). We are able to shift our attention independent of eye movements, yet eye movements are functionally coupled to shifts in attention (Deubel and Schneider, 1996; Henderson, 1993). The physical structure of our eyes is such that we can only process colour information in the central areas of our retina, while those areas in the periphery are in essence colour-blind (Coren et al., 1993). Therefore, in order to process colour information, it is necessary to focus on the target stimuli. The exact role of eye movements in the integration of colour information is not known, yet there is no question that we must first fixate on this information to perceive it.

Eye movements are typically an indication of the viewer’s spatial focus of attention on a display (Gefen et al., 2003). Eyes naturally fixate on areas that are surprising, salient or important related to previous experience (Lavie and Tractinsky, 2004). Each website used in this study was divided into Areas of Interest (AoI). Participants were permitted to view the web pages as long as desired but in alignment with Pan et al. (2004) only the first 15 seconds of viewing were analyzed. In this study, minimum duration time for a fixation is .05 seconds (following Lankförd, 2000) and is expected to represent interest in the viewed portion of the website. Gazetracker software was used to process ocular data.

The eye-tracker system used was Applied Science Laboratories Model 504 with head tracking integration. Eye movements were processed using a small camera mounted on a pan/tilt optics mechanism positioned under the stimulus monitor. Participants wore a headband with a small mounted sensor, allowing the pan/tilt mechanism to track head movements without loss of eye image. This permitted participants to move their heads in a relatively natural manner. Each participant was screened to ensure normal or corrected vision, and those wearing hard contact lenses or eye glasses were not considered for the experiment. Suitable participants were calibrated with the eye-tracking equipment by looking at an image of nine numbered dots on the screen. The researcher called out a number and participants were instructed to look at each dot as called. Any calibration errors were immediately corrected. Participant calibration with the eye tracker took about five to ten minutes. Following calibration, participants were presented with the experimental websites with instructions to examine each site as they would normally, with the intention to gather information for product and
company assessment. Each participant’s gaze was monitored during the eye-tracking procedure to ensure accurate calibration throughout the testing phases. If calibration problems were detected then the participant was recalibrated.

Eye-tracking analysis can proceed in either a top-down or bottom-up fashion. Top-down analysis is based on theoretical hypotheses, whereas a bottom-up approach is based entirely on observation of the data without predefined theories (Gefen et al., 2000). Our hypotheses are derived from extant literature. As such, the eye-tracking analysis of this research follows a top-down approach.

4.3.2. Survey and instrument validation

A survey was administered after each participant completed the assigned task for each website condition. All items in the survey are constructed as agree-disagree statements on a five-point Likert scale. The survey appears in Appendix C.

Content validity considers how representative and comprehensive the items are in creating the experimental constructs. It is assessed by examining the process by which the construct items are generated (Short et al., 1976). Constructs should draw representative questions (items) from a universal pool (Corritore et al., 2003). In this research, survey items are adapted from previously validated work on satisfaction (Cyr, 2008), trust3 (Gefen, 2000; Yoon, 2002), and e-loyalty (Cyr, 2008). Therefore, content validity for these three constructs is established through literature review (Short et al., 1976). The colour appeal scale is new and could not be established through literature review because expert judges are employed to evaluate the construct (as per Straub, 1989). Two usability experts independently listed attributes relevant to the appeal of colour on an e-commerce website. The lists were compared for commonality, resulting in seven salient attributes which were used as the basis for the new colour appeal construct. The usability experts agreed the resulting seven items were reasonable and appropriate for measuring website colour appeal. The entire instrument was pre-tested with a small sample of five participants who were asked to provide detailed comments on any wording or concept confusion. Based on this feedback, some minor wording adjustments were made before conducting the full study.

Construct validity assesses the extent to which a construct measures the variable of interest and whether “the measures chosen ‘fit’ together in such a way as to capture the essence of the construct” (Schmitt and Pan, 1994, p. 388). There should be high correlations between items of the same construct (convergent validity), and low correlations between items of different constructs (discriminant validity) (Short et al., 1976). To assess convergent validity of the measurements Fornell and Larcker (1981) proposed examining three metrics: (i) the item reliability of each measure; (ii) the composite (construct) reliability of each construct; and (iii) the average variance extracted for each construct. Item reliability of each measure is assessed by performing a principle components factor analysis (PCA) as recommended by Straub (1989). A construct exhibits satisfactory convergent and discriminant validity when items load highly on their related factor and have low loadings on unrelated factors. Items that do not load properly may be dropped from the instrument (Churchill, 1979). As per Hair et al. (1995), CA-6, CA-7, and S-4 were removed due to low loadings on their respective constructs and/or high cross-loadings on other constructs. The final loadings and cross-loadings matrix for the entire sample (n = 270) is shown in Table 3.

In addition, convergent and discriminant validity is demonstrated at the individual culture level. Table 4 summarizes the factor loadings for each culture.

While there are some variances in item loadings across cultures they are slight, and all items load appropriately on their respective constructs. Therefore, the survey instrument demonstrates item reliability of each measure at the aggregate level as well as the individual culture level.

Construct reliability is assessed using Cronbach’s \( \alpha \)-values. As shown in Table 4, \( \alpha \)-values ranged from .877 (for satisfaction) to .934 (for e-loyalty). Rivard and Huff (1988) suggested this measure for reliability should be higher than .5 and ideally higher than .7. Nunnally (1978) also recommends the Cronbach’s \( \alpha \) of a scale should be greater than .7 for items to be used together as a construct.

Table 3

<table>
<thead>
<tr>
<th>Component</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA1</td>
<td>.861</td>
<td>.192</td>
<td>.280</td>
<td>.103</td>
</tr>
<tr>
<td>CA2</td>
<td>.854</td>
<td>.170</td>
<td>.244</td>
<td>.159</td>
</tr>
<tr>
<td>CA3</td>
<td>.734</td>
<td>.198</td>
<td>.247</td>
<td>.142</td>
</tr>
<tr>
<td>CA4</td>
<td>.825</td>
<td>.051</td>
<td>.181</td>
<td>.158</td>
</tr>
<tr>
<td>CA5</td>
<td>.840</td>
<td>.077</td>
<td>.084</td>
<td>.110</td>
</tr>
<tr>
<td>CA1</td>
<td>.248</td>
<td>.360</td>
<td>.180</td>
<td>.150</td>
</tr>
<tr>
<td>T1</td>
<td>.152</td>
<td>.369</td>
<td>.188</td>
<td>.213</td>
</tr>
<tr>
<td>T2</td>
<td>.061</td>
<td>.844</td>
<td>.239</td>
<td>.204</td>
</tr>
<tr>
<td>T3</td>
<td>.199</td>
<td>.252</td>
<td>.259</td>
<td>.855</td>
</tr>
<tr>
<td>S2</td>
<td>.236</td>
<td>.301</td>
<td>.311</td>
<td>.795</td>
</tr>
<tr>
<td>L1</td>
<td>.228</td>
<td>.270</td>
<td>.769</td>
<td>.387</td>
</tr>
<tr>
<td>L2</td>
<td>.427</td>
<td>.241</td>
<td>.794</td>
<td>.177</td>
</tr>
<tr>
<td>L3</td>
<td>.335</td>
<td>.310</td>
<td>.794</td>
<td>.272</td>
</tr>
</tbody>
</table>

\( \alpha \)-value: .921 .901 .934 .877

AVE: .679 .736 .617 .682

---

3It is recognized a multi-dimensional construct for trust is appropriate in research when online trust is the primary focus. Trust may result from a consumer’s belief that an online vendor demonstrates ability, benevolence or integrity (Loftus and Mackworth, 1978). Alternately, in studies such as this one when trust is one element included to better understand a more comprehensive user reaction to a website, a single construct for trust has been used (Gefen, 2000; Kay et al., 1991).
Table 4
Factor loadings for the full sample and individual cultures.

<table>
<thead>
<tr>
<th></th>
<th>Full sample (n=270)</th>
<th>Canada (n=90)</th>
<th>Germany (n=90)</th>
<th>Japan (n=90)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA1</td>
<td>.861</td>
<td>.727</td>
<td>.869</td>
<td>.819</td>
</tr>
<tr>
<td>CA2</td>
<td>.854</td>
<td>.770</td>
<td>.793</td>
<td>.874</td>
</tr>
<tr>
<td>CA3</td>
<td>.734</td>
<td>.576</td>
<td>.680</td>
<td>.846</td>
</tr>
<tr>
<td>CA4</td>
<td>.825</td>
<td>.664</td>
<td>.876</td>
<td>.773</td>
</tr>
<tr>
<td>CA5</td>
<td>.840</td>
<td>.863</td>
<td>.888</td>
<td>.863</td>
</tr>
<tr>
<td>T1</td>
<td>.860</td>
<td>.844</td>
<td>.861</td>
<td>.828</td>
</tr>
<tr>
<td>T2</td>
<td>.869</td>
<td>.855</td>
<td>.852</td>
<td>.900</td>
</tr>
<tr>
<td>T3</td>
<td>.844</td>
<td>.857</td>
<td>.845</td>
<td>.876</td>
</tr>
<tr>
<td>S2</td>
<td>.855</td>
<td>.821</td>
<td>.710</td>
<td>.879</td>
</tr>
<tr>
<td>S3</td>
<td>.795</td>
<td>.764</td>
<td>.671</td>
<td>.850</td>
</tr>
<tr>
<td>L1</td>
<td>.769</td>
<td>.775</td>
<td>.841</td>
<td>.696</td>
</tr>
<tr>
<td>L2</td>
<td>.794</td>
<td>.766</td>
<td>.833</td>
<td>.811</td>
</tr>
<tr>
<td>L3</td>
<td>.794</td>
<td>.818</td>
<td>.771</td>
<td>.830</td>
</tr>
</tbody>
</table>

Table 5
Discriminant validity.

<table>
<thead>
<tr>
<th></th>
<th>Colour appeal</th>
<th>Trust</th>
<th>Satisfaction</th>
<th>Loyalty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colour appeal</td>
<td>.824</td>
<td>.403</td>
<td>.477</td>
<td>.630</td>
</tr>
<tr>
<td>Trust</td>
<td>.403</td>
<td>.858</td>
<td>.563</td>
<td>.581</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>.477</td>
<td>.563</td>
<td>.826</td>
<td>.682</td>
</tr>
<tr>
<td>Loyalty</td>
<td>.630</td>
<td>.581</td>
<td>.682</td>
<td>.786</td>
</tr>
</tbody>
</table>

The diagonal elements in bold (the square root of the average variance extracted) should exceed the inter-construct correlations below and across them for adequate discriminant validity.

Therefore, all the constructs in this study demonstrate construct reliability. Fornell and Larcker (1981) suggested the average variance extracted from a construct should exceed .5. As shown in Table 4, all constructs exceed this criterion. Thus, the proposed constructs demonstrate convergent validity on all three metrics proposed by Fornell and Larcker (1981).

Discriminant validity is assessed to ensure the constructs differ from each other. As per Fornell and Larcker (1981) the correlations between items in any two constructs should be lower than the square root of the average variance shared by items within a construct. As shown in Table 5, the square root of the variance shared between a construct and its items is greater than the correlations between items in any two constructs differ from each other. Therefore, the survey demonstrates satisfactory content validity (established through literature reviews and expert judges); satisfactory convergent validity (as evidenced from the PCA analysis, \( \alpha \)-values, and AVE values); and satisfactory discriminant validity (as evidenced from inter-construct correlation analysis).

4.3.3. Interviews

Interview questions followed the survey and participant responses were recorded using a digital recorder. Atlas.ti software provides an effective means to analyze qualitative data such as interview transcripts, and is used for content analysis and coding. The analysis process consists of the following steps: (1) data preparation (i.e. interview transcription and formatting); (2) in vivo coding (use of participants’ words as code labels) and open coding (use of arbitrary labels for code labels); (3) category and concept building in which semantic relationships between codes are identified to build higher conceptual abstractions; and finally (4) theory building based on interpretation of the results. While it is customary to have multiple raters code the same data when only one (qualitative) methodology is used, in an instance where multiple data sources are used to confirm the same phenomenon inter-rater reliability is not critical. Armstrong et al. (1997) note that when triangulation of data is used (as in the current research), then use of these multiple methodologies lends weight to the findings.

5. Results

5.1. Impact of colour appeal on trust (H1) and satisfaction (H2) and impact of trust and satisfaction on e-loyalty (H3 and H4)

A structural equation modeling (SEM) approach is adopted to test H1 through H4. SEM simultaneously tests structural and measurement models (Armstrong et al., 1997) and provides a more complete analysis for inter-relationships in a model (Flavián et al., 2006). The variance-based PLS method was chosen over covariance-based methods, such as LISREL. PLS is relatively robust to deviations from a multivariate distribution (Gage, 1999) and supports exploratory and confirmatory research (Gage, 1999). Since PLS does not generate an overall goodness-of-fit index (as with LISREL), model validity is primarily assessed by examining structural paths and \( R^2 \) values (Churchill, 1979). As recommended by Chin (1998), bootstrapping (with 500 sub-samples) was performed to test the statistical significance of each path coefficient using \( t \)-tests. The PLS model is shown in Fig. 2.

All path coefficients of the hypothesized causal links are significant (\( t \)-values shown in parentheses under the path beta values). Thus, H1 through H4 are supported. Approximately 52% of the variance in the e-loyalty construct is accounted for by the variables in the model (\( R^2= .552 \)). All \( R^2 \) values of the endogenous constructs in the model exceed the 10% benchmark recommended by Falk and Miller (1992).

Chin (1998) and Gefen et al. (2000) advise the minimum sample size for a PLS analysis should be the larger of (i) 10 times the number of items for the most complex construct; or (ii) 10 times the largest number of independent variables impacting a dependent variable. In our model, the most complex construct (colour appeal) has five items and the largest number of independent variables estimated for a dependent variable is only two (for e-loyalty). The total sample size for this study is 270 which is more than adequate for PLS estimation procedures.
5.2. Culture as a moderator between colour appeal and trust (H5a), and between colour appeal and satisfaction (H5b)

Separate PLS models were run for each country. Table 6 summarizes PLS results comparing the full sample with individual country samples where colour appeal is the exogenous (independent) variable.

Causal paths between colour appeal and trust and between colour appeal and satisfaction are significant across all culture groups. Carte and Russell (2003) stress that investigators should examine the change in $R^2$ values when drawing conclusions about relative moderator effect sizes. As shown in Table 6, there are some differences in $R^2$ values across country groups (for example $R^2 = .228$ for trust in the Canadian sample but is only .095 in the Japanese sample). This observation prompted further analysis.

Table 7

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Independent variable</th>
<th>Path coefficient</th>
<th>$t$-Value</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trust</td>
<td>Colour appeal (CA)</td>
<td>.274</td>
<td>2.950**</td>
<td>.205</td>
</tr>
<tr>
<td></td>
<td>Culture</td>
<td>.053</td>
<td>.562</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CA × Culture</td>
<td>.207</td>
<td>1.802</td>
<td></td>
</tr>
<tr>
<td>Satisfaction</td>
<td>Colour appeal (CA)</td>
<td>.277</td>
<td>2.173*</td>
<td>.316</td>
</tr>
<tr>
<td></td>
<td>Culture</td>
<td>.115</td>
<td>.872</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CA × Culture</td>
<td>.272</td>
<td>1.629</td>
<td>1.629</td>
</tr>
</tbody>
</table>

** $p$-value < .01.

* $p$-value < .05.

Culture was coded using two 0/1 dummy variables to categorically capture the three countries. Model results, including moderators, are shown in Table 7. A partial least squares product indicator approach for measuring interaction is employed as suggested by Chin et al. (1996). Although the beta values are encouraging, interactions fall just short of significance ($t = 1.802$ when trust is the dependent variable and $t = 1.629$ when satisfaction is the dependent variable). Therefore, based on the sample in this research, the influence of colour appeal on trust and satisfaction is not statistically moderated by culture.

5.3. Impact of colour schemes on culture groups for colour appeal (H6a), trust (H6b), and satisfaction (H6c)

Survey, eye-tracking, and interview data were analyzed and compared to explore the impact of colour schemes on the various cultural groups. First, survey data was analyzed to determine if colour schemes elicit different reactions within cultural groups. Table 8 provides mean scores for colour appeal, trust, and satisfaction for the three colour conditions across three countries (Canada, Germany, and Japan). Table 9 illustrates pair-wise colour comparisons by country.

Survey analysis reveals there are some significant differences in perceptions across the colour conditions for
the three cultures. In terms of colour appeal, it is interesting to note that all three cultures tend to dislike the yellow websites. Canadians found both the grey and blue colour schemes more appealing than the yellow ($p < .01$ and $p < .001$ respectively). Similarly, appeal for the yellow website was significantly lower than the blue website for Germans and Japanese ($p < .01$, and $p < .05$ respectively). Contrary to predictions, Japanese did not like this brighter tone. Canadians and Germans even showed distrust for the yellow colour scheme when compared to the blue colour scheme ($p < .05$ for both). In terms of preferences, as predicted survey results reveal a pronounced liking of the blue colour scheme among the German participants.

Exploring further, Table 10 illustrates pair-wise country comparisons by colour scheme. For the grey colour scheme, Canadians found the colour more appealing than Germans ($p < .001$) and the website more satisfying than did Germans ($p < .001$). This grey website was also more satisfying for Canadians as compared to Japanese ($p < .01$). For the blue colour scheme Germans trusted the site more than Japanese ($p < .05$), and Canadians were more satisfied with the website than Japanese ($p < .01$). Therefore, as expected survey results indicate Canadians were more favorable in their perceptions toward the grey colour scheme, when compared to German and Japanese groups. Japanese, on the other hand, seemed to be less favorable in their perceptions toward the blue colour scheme, when compared to Canadians and Germans. There were no significant differences in perceptions of the yellow colour scheme across cultures, as all cultures have a general aversion toward this scheme.

Second, eye-tracking data was analyzed for the three colour schemes across cultures. For each version of the experimental website two primary zones (left navigation bar and top graphic) were manipulated to

Table 9
Multiple comparisons for colour appeal, trust and satisfaction for the three colour scheme conditions by cultures.

<table>
<thead>
<tr>
<th>Country</th>
<th>Colour scheme paired comparison</th>
<th>Colour appeal (Sig.)</th>
<th>Trust (Sig.)</th>
<th>Satisfaction (Sig.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canada</td>
<td>Grey vs. Blue</td>
<td>.175</td>
<td>.925</td>
<td>.996</td>
</tr>
<tr>
<td></td>
<td>Grey vs. Yellow</td>
<td>.005**</td>
<td>.065</td>
<td>.067</td>
</tr>
<tr>
<td></td>
<td>Blue vs. Yellow</td>
<td>.000***</td>
<td>.026*</td>
<td>.055</td>
</tr>
<tr>
<td>Germany</td>
<td>Grey vs. Blue</td>
<td>.000***</td>
<td>.126</td>
<td>.027*</td>
</tr>
<tr>
<td></td>
<td>Grey vs. Yellow</td>
<td>.077</td>
<td>.727</td>
<td>.804</td>
</tr>
<tr>
<td></td>
<td>Blue vs. Yellow</td>
<td>.005**</td>
<td>.021*</td>
<td>.118</td>
</tr>
<tr>
<td>Japan</td>
<td>Grey vs. Blue</td>
<td>.008**</td>
<td>.774</td>
<td>.421</td>
</tr>
<tr>
<td></td>
<td>Grey vs. Yellow</td>
<td>.769</td>
<td>.395</td>
<td>.940</td>
</tr>
<tr>
<td></td>
<td>Blue vs. Yellow</td>
<td>.049*</td>
<td>.121</td>
<td>.252</td>
</tr>
</tbody>
</table>

Notes: The Tukey test was used for the multiple comparisons.

***$p$-value < .001.
** $p$-value < .01.
* $p$-value < .05.

Table 10
Multiple comparisons for colour appeal, trust and satisfaction for the three cultures by colour scheme conditions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Country paired comparison</th>
<th>Colour appeal (Sig.)</th>
<th>Trust (Sig.)</th>
<th>Satisfaction (Sig.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grey colour scheme</td>
<td>Japan vs. Germany</td>
<td>.145</td>
<td>.660</td>
<td>.840</td>
</tr>
<tr>
<td></td>
<td>Japan vs. Canada</td>
<td>.105</td>
<td>.096</td>
<td>.001**</td>
</tr>
<tr>
<td></td>
<td>Germany vs. Canada</td>
<td>.000***</td>
<td>.440</td>
<td>.000***</td>
</tr>
<tr>
<td>Blue colour scheme</td>
<td>Japan vs. Germany</td>
<td>.666</td>
<td>.019*</td>
<td>.208</td>
</tr>
<tr>
<td></td>
<td>Japan vs. Canada</td>
<td>.913</td>
<td>.114</td>
<td>.009**</td>
</tr>
<tr>
<td></td>
<td>Germany vs. Canada</td>
<td>.893</td>
<td>.736</td>
<td>.379</td>
</tr>
<tr>
<td>Yellow colour scheme</td>
<td>Japan vs. Germany</td>
<td>.993</td>
<td>.520</td>
<td>.883</td>
</tr>
<tr>
<td></td>
<td>Japan vs. Canada</td>
<td>.242</td>
<td>.547</td>
<td>.094</td>
</tr>
<tr>
<td></td>
<td>Germany vs. Canada</td>
<td>.293</td>
<td>.999</td>
<td>.238</td>
</tr>
</tbody>
</table>

Notes: The Tukey test was used for the multiple comparisons.

***$p$-value < .001.
** $p$-value < .01.
* $p$-value < .05.
reflect the colour scheme conditions. Time spent viewing these colour zones, as a percent of the total time spent viewing the web page, is graphically displayed in Fig. 3.

There were no significant differences in the proportion of time spent viewing the colour manipulation zones across different cultures. Although the yellow manipulations are most disliked across cultures, on average participants did not spend less gazing time within these zones.

Refining the eye-tracking analysis to compare across cultures reveals some interesting results. As shown in Fig. 4, Canadian and Japanese participants did not show great variations in the proportion of time spent gazing in the manipulated zones across colour schemes. However, Germans spent approximately 31% of website viewing time in the yellow zones, but only 19% of website viewing time in the corresponding blue zones. Recall that Germans demonstrated greatest preference for the blue colour scheme and a strong aversion to the yellow colour scheme (as per the survey results). One would expect viewing time to be higher in zones that are more appealing rather than zones that are less appealing. Our eye-tracking analysis discovered the opposite for the German group. Very similar results are revealed when analyzing the proportion of total fixations. These unexpected results are explored further when triangulated with the qualitative interview data.

Third, interview data is analyzed to better understand website colour scheme preferences across cultures. As previously noted, interview data is coded using two methods: (i) in vivo (using the participants’ exact words as the basis for a code), and (ii) open coding (using arbitrary labels to code the data). Categories are then developed to identify relationships between codes, followed by the creation of more theoretical entities called concepts. Five main concepts about website colour emerge from the data:

**Aesthetics:** Participant perceptions that the website colour provides artistic appeal. Words such as “beautiful”, “ugly”, “earthy”, “cartoonish”, and “jarring” encapsulate the idea of aesthetics.

**Affective:** Affective properties are the emotional qualities of the manipulated colours. The codes “friendly”, “serious”, “fun”, and “dull” are examples of affective properties.

**Functional:** Effects of colour on perceived structural, organizational, and usable qualities of the website. Examples include “easy to read”, “distracting”, and “confusing”.

**Harmony:** Balance between manipulated colours with the rest of the website. Harmony is represented by words such as “well balanced”, “good colour combinations”, and “matched other colours”.

**Appropriateness:** Given that the chosen website was designed for a particular purpose (i.e. to sell electronics), participants made comments regarding how appropriate the colours were for this type of website. Comments such as “appropriate”, “trying too hard”, “bad colour for product presentation”, and “the colour of products should be the same as site” reflect this concept.

Table 11 contains a summary of the above emergent concepts expressed during the interviews, across colour schemes and cultures. In some cases, participants from a culture mentioned that a particular website had positive colour qualities related to the above identified concepts, whereas other participants from the same culture expressed negative reactions to the same concept. For instance, some Canadian participants felt the blue colour scheme had positive affective qualities (i.e. “friendly”) and others felt the site’s affective qualities were negative (i.e. “too neutral”). While Canadian and Japanese participants made reference to all of the above emergent categories in their interviews, it is interesting that no Germans made explicit reference to aesthetics as having either a positive or negative valence for any of the websites.

In general, interview analysis supports quantitative results and provides some possible insights into the eye-tracking results. Blue and grey colour schemes were preferred over yellow colour scheme across cultures. While there were some negative affective comments (mostly focused on “dullness”) for the blue scheme, this colour
did not elicit negative comments in terms of aesthetics, appropriateness or harmony. In contrast, the yellow colour scheme elicited negative comments for all the emerging concepts. In particular, Germans were most vocal in terms of the poor functionality, disharmony and inappropriateness of the yellow website. German participants noted the yellow colour scheme is “showy” or “too friendly”. They further note the yellow is: “[A] bad colour for product presentation”, “distracting”, “should only be used to highlight important things”, and “needs to be combined with other colours”. Such sentiments could account for the increased draw of visual attention as demonstrated in the eye-tracking analysis for this culture. As noted earlier, eye fixation is required to process colour information. When colour is distracting and at odds with cultural expectations, it is reasonable to observe increased visual attention as viewers try to overcome the confusion caused by the unusual visual. Earlier work on the use of human images in website design (Cyr et al., 2006) found similar results, where eye fixations were significantly higher in conditions that were unexpected and distracting.

5.4. Summary of results

The proposed research model was tested for relationships between website colour appeal on antecedents of online loyalty (e-loyalty), namely online trust and satisfaction. Potential moderating effects of culture on these relationships were also investigated. Online trust and satisfaction were confirmed as strong predictors of e-loyalty (H3 and H4) in a context of colour, and the newly developed colour appeal construct was shown to significantly influence online trust and satisfaction (H1 and H2). Although these results appear encouraging, we can not conclude (statistically) that culture moderates the relationship between colour appeal and trust (H5a), and between colour appeal and satisfaction (H5b).

Further, the model tested if cultural groups vary in their reaction to various website colour schemes (H6). Both similarities and differences were found among cultures. All cultures tended to dislike the yellow colour scheme websites. Germans had the most pronounced preference for the blue colour scheme and Canadians appreciated the grey colour scheme more than Germans and Japanese. Interview and eye-tracking analysis supported and expanded quantitative survey results. Table 12 is a summary of the hypotheses tested as well as a summary of key results.

6. Discussion and conclusions

Although various elements of website design have significant impacts on consumer perceptions and willingness to interact with or return to a website (Fornell and Larcker, 1981; Szymanski and Hise, 2000), research is sparse when examining potential impacts of website colour (Marcus and Gould, 2000). Especially rare are rigorous studies of colour in website design across cultures (Meyers-Levy and Peracchio, 1995). The current investigation provides an exploration of this under-examined domain.

A research model with a multiple focus was tested drawing on literature from human computer interaction, communications, and marketing. The model encompasses well established e-commerce variables of trust, satisfaction and e-loyalty. A new colour appeal construct was introduced to provide additional insights into how online consumers perceive the colour attributes of websites. A culture dimension is also incorporated in the research model to explore moderating effects on consumer perceptions. To verify the proposed model multiple research methods are employed including surveys, interviews, and eye-tracking.

From a theoretical perspective, this investigation provides the following contributions:

1. Trust and satisfaction are confirmed as antecedents to e-loyalty within a new context. While these relationships have been supported by others (Falk and Miller, 1992; Kim and Benbasat, 2006; Lankford, 2000) this was the first investigation in which these relationships are validated with various website colour treatment manipulations.
Table 12
Summary of Hypotheses and results using three methodologies.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Findings</th>
<th>Survey data</th>
<th>Eye-tracking</th>
<th>Interviews</th>
</tr>
</thead>
</table>
| H1: For all cultural groups, increased colour appeal will result in greater trust | Supported | $t = 7.89, p < .001$ | For pooled data, no significant differences in the proportion of time spent viewing the colour manipulation zones. | Five concepts emerged:  
- Aesthetic  
- Affective  
- Functional  
- Harmony  
- Appropriateness |
| H2: For all cultural groups, increased colour appeal will result in greater satisfaction | Supported | $t = 9.69, p < .001$ | Germans spent approximately 31% of their website viewing time in the yellow zones but only 19% of their website viewing time in the corresponding blue zones. | Germans made no explicit reference to aesthetics as neither a positive or negative for any of the websites. |
| H3: For all cultural groups, higher trust will result in greater e-loyalty | Supported | $t = 4.91, p < .001$ | | Blue scheme did not elicit negative comments in terms of its aesthetics, appropriateness or harmony. Yellow colour scheme elicited negative comments for all the emerging concepts. Germans were most vocal in terms of the poor functionality, harmony and appropriateness of the yellow website. |
| H4: For all cultural groups, increased satisfaction will result in greater e-loyalty | Supported | $t = 9.53, p < .001$ | | |
| H5a: The influence of colour appeal on trust will be moderated by culture | Not supported | $t = 1.802, p > .05$ | | |
| H5b: The influence of colour appeal on satisfaction will be moderated by culture | Not supported | $t = 1.629, p > .05$ | | |
| H6a: Most appealing colours will be yellow for Japanese, blue for Germans, and grey for Canadians. | Supported for Canadians | Grey: more appealing for C (p < .001) and over J (p < .01) | | |
| H6b: Most trusted colours will be yellow for Japanese, blue for Germans, and grey for Canadians. | Supported for Germans | Blue: more trusted for G over J (p < .05) | | |
| H6c: Most satisfying colours will be yellow for Japanese, blue for Germans, and grey for Canadians. | Not supported | Yellow: significantly disliked across cultures | | |

2. A new colour appeal scale is validated with potential for application in future research.

3. Website colour appeal is shown to be a significant determinant for both website trust and satisfaction. This is a new finding for the literature and extends work by Karvonen (2000) on website aesthetics and trust, Latomia and Happ (1987) related to the ineffective use of colour impeding consumer satisfaction, and Kim and Moon (1998) with respect to colour and trust. The current research also extends earlier theoretical contributions regarding cognitive-affective communication and the theory of theoretical reversals now applied to colour treatments in a website context with subsequent consequences of trust and satisfaction.

4. Culture research has been advanced through demonstration of how responses and reactions to website colour treatments differ among Canadian, German and Japanese cultures as tested in a controlled laboratory setting. According to Noiwan and Norcio (2006) prior studies of colour in website design across cultures have tended to lack the rigor present in the current research.

5. Triangulation of data provides rich insights in the current investigation, and demonstrates the value of multiple methodologies in future research. In particular, this investigation provides an illustrative example of how the often misunderstood eye-tracking technique can complement other research methodologies.
From a practical perspective there are numerous contributions that emanate from this work. In particular, designers of websites and web advertisements are advised to pay attention to colours employed in website design. Colour has the potential to communicate meaning to the user and influence perception of the website. The appeal an Internet user feels toward a particular colour scheme can have a significant impact on his or her experience, with implications for trust, satisfaction, or future interaction with the website. In the online world where consumer confidence is often low and fragile, subtle design differences can push online users toward either completing transactions or abandoning shopping carts. For example, in this study a yellow colour scheme received adverse reaction from most participants when utilized on an electronics e-commerce site. This is counter to Boor and Russo (1993) who wrote that Japanese associate yellow with nobility and grace. Not only is yellow perceived as unappealing in this context, but it also caused distrust among participants. Alternately, a blue colour scheme generally was viewed favorably for an electronics e-commerce site. This is in line with other research in which blue is universally liked (Nielsen and Del Galdo, 1996) and generally associated with “wealth, trust, and security” (Liichtle, 2007).

Further, designers and others involved in e-commerce should consider the impact of colour design related to culture. Websites that are localized to reflect the preferences and content unique to the culture have the potential to increase “stickiness” and e-loyalty (Schijns, 2003; Singh et al., 2003). Those designers who better understand the preferences for their target online audience are more likely to achieve success in highly competitive online markets. For example, in this study Canadians have a stronger preference for a grey colour scheme for an electronics e-commerce site when compared to Germans and Japanese. Germans, on the other hand, show a stronger preference for a blue colour scheme for this type of website and appear more sensitive to jarring, unnatural or unappealing colours. These findings parallel those of Cyr and Trevor-Smith (2004). In fact for Germans, prolonged exposure to colours perceived as unpleasant result in ‘hostility’ toward the website and waning trust. These differential experiences by members of different country groups serve to support earlier work that website elements vary by culture (Bhattacherjee, 2002; Cyr et al., 2006; Cyr et al., 2005; Fornell, and Larcker, 1981; Meyers-Levy and Peracchio, 1995), although now tested using multiple methods in a rigorous experimental setting.

While the results tend to be in the expected direction, they are not as strong as anticipated. Speculating as to why this occurred, it may be that colour has universal components as posited by Kay et al., (1991) which serve to neutralize some of our findings. As a counterargument to website localization in which specific preferences of cultural groups are accommodated, it may be that certain commonalities exist among cultures that trump local socio-cultural affective meaning (Adams and Osgood, 1973).

According to these authors, one source of commonality among people is exposure to similar colours in the environment (e.g., green forests, blue skies, yellow sun). Regardless to which theory one subscribes, there appears to be an emotional component of user reactions to colour in the context of websites. It would be of further interest to examine to what extent website localization through colour schemes has relative importance compared to other design characteristics such as use of images which is also known to have a hedonic or emotional impact on users (Fogg et al., 2002; Zheng et al., 2002).

Application of design concept typologies (i.e. aesthetics, affective, functional, harmony, appropriateness) may also be valuable to web developers. When explaining user reactions toward a website, Canadians and Japanese touched upon all five typology dimensions when assessing positive and negative elements. However, Germans do not mention any words or phrases that pertained to either positive or negative aesthetics of the experimental websites and may be indicative of differing design priorities among cultures.

As with all empirical studies, there are certain limitations to this investigation which may provide an impetus for future research. First, while the experimental setting is required due to the nature of the data collection process, it presents an artificial environment for online shopping. Additional research can compare results obtained in settings that are more natural to the user.

Second, the number of participants in this study is not large due to the in-depth nature of the data collection process. A student population was recruited from international business schools. Each group is representative of the home culture, but participants are generally younger than the general population. Further, only three cultures are compared and generalizations should be cautioned beyond these specific cultures. Future research can be focused on larger sample sizes and a broader base of Web users. However it is important to note that a student sample is representative and appropriate for e-retailing research as students are frequent users of the Internet for communication and commercial transactions (Tian and Emery, 2002; Walczuch and Lundgren, 2004).

Third, one website is chosen and experimentally manipulated for three colour treatments. The advantage of this procedure is that the websites and colours observed by the user are highly controlled. However, as was noted by participants, some colour manipulations appear unnatural for the electronics e-commerce website context. Additional research can therefore explore differing colours within this context as well as in diverse contexts.

Fourth, the exact nature of the relationship between attention, eye movements, colour perception, and higher-level psychological reactions (e.g., emotion and behavior) is not yet known. The assumption in the present study was that eye fixations measured participants’ interest on the fixated area, whether drawn from a pleasing or distracting source. Yet, websites are highly complex multi-dimensional
stimuli, which could have influenced participants’ ocular responses at a more fine-grained level of detail. For example, previous research has shown that aside from hue, brightness and saturation may also influence a person’s emotional reactions to colours (Valdez and Mehrabian, 1994). Given the ability of the eye tracker to capture detailed eye movement behavior, taking account of these and other variables such as display luminosity is suggested for future research.

To conclude, the impact of varying website colour schemes on online users is demonstrated, and supports that colour has the potential to influence user perceptions, emotions, reactions, and behavioral intentions (Latomia and Happ, 1987). As the world becomes increasingly global, cultural considerations are central for website design. Hence, it is hoped this research is a catalyst for future studies to expand theoretical understanding at the intersection of design and culture. This avenue of research further has potential to provide Web designers with enhanced knowledge for how to attract e-loyal consumers through the effective use of colour.

Acknowledgement

Funding for the research is provided by the Social Sciences and Humanities Research Council of Canada.

Appendix A. Considerations for pooled data

Overall sample size for the study is 270 (30 participants by 3 countries by 3 colour treatments). Some of the analysis in this paper is conducted on pooled repeated measures data. West and Hepworth (1991) outlined the most critical concern when pooling repeated measures from the same subjects is the possibility of correlated errors. Therefore, to appropriately pool data the error terms should be uncorrelated. As per Venkatesh et al. (2003) the endogenous variable (e-loyalty in this study) was examined for error term correlations. Error terms associated with the e-loyalty construct are calculated for grey, blue, and yellow colour scheme manipulations for each culture separately, and across cultures. The above table shows error term correlations across the e-loyalty construct for colour manipulations (which were repeated measures for each subject). All error correlations are insignificant and therefore the pooling of data across colour scheme manipulations is appropriate (see Table A1).

Appendix B. Sample screenshots

These images depict web pages for Canada (top), Germany (middle), and Japan (bottom). Regions on the left side and top of the page (in bold black frames) received the colour treatments (denoted as ‘colour zones’).
Appendix C. Participant survey

See Table C1.

References


