Association for Information Systems
AIS Electronic Library (AISeL)

ICIS 2010 Proceedings

International Conference on Information Systems (ICIS)

2010

COLORING IN THE LINES: USING COLOR TO CHANGE THE PERCEPTION OF QUALITY IN E-COMMERCE SITES

Christopher M. Conway
University of Arkansas, cmconway@uark.edu

Jean-Eric Pelet
University of Nantes, jepelet@yahoo.com

Panagiota Papadopoulou
University of Athens, peggy@di.uoa.gr

Moez Limayem
University of Arkansas, cmconway@uark.edu

Follow this and additional works at: http://aisel.aisnet.org/icis2010_submissions

Recommended Citation
Conway, Christopher M.; Pelet, Jean-Eric; Papadopoulou, Panagiota; and Limayem, Moez, "COLORING IN THE LINES: USING COLOR TO CHANGE THE PERCEPTION OF QUALITY IN E-COMMERCE SITES" (2010). *ICIS 2010 Proceedings*. 224.
http://aisel.aisnet.org/icis2010_submissions/224

This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICIS 2010 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

COLORING IN THE LINES: USING COLOR TO CHANGE THE PERCEPTION OF QUALITY IN E-COMMERCE SITES

Completed Research Paper

Christopher M. Conway
University of Arkansas
204 Business Building
Fayetteville, AR 72701
cmconway@uark.edu

Jean-Eric Pelet
University of Nantes
16 rue Laurence Savart
F – 75020 Paris
jepelet@yahoo.com

Panagiota Papadopoulou
University of Athens
Department of Informatics and Telecommunications
Panepistimiopolis, GR - 15784, Athens, Greece
peggy@di.uoa.gr

Moez Limayem
University of Arkansas
204 Business Building
Fayetteville, AR 72701
cmconway@uark.edu

Abstract

E-commerce sites, like their brick-and-mortar cousins, are designed to draw in potential customers and encourage purchases and brand loyalty. Currently, E-commerce website designs are often rejected by their user community, which costs money through wasted development and lost customers. User acceptance is the goal of every design choice – and yet little work has extrapolated from the known affective responses users have to color to website design. We posit that a better understanding of user reactions to color will increase user perceptions of quality. We hypothesize that color saturation and brightness cause an affective reaction in the user, which can change the user's trust in the site, and through this their perception of a site's quality. In an experimental lab study, we find support for our hypotheses, providing suggestions for future website design to improve customer acceptance.

Keywords: Color, Trust, Emotion, E-Commerce, Human-Computer Interaction

Introduction

I'm feeling blue today.

We've all heard someone say this. We know that they are not saying that the light reflecting off them is predominately around 450 nanometers, but that they feel sad or depressed. Other colors have similar connotations; we might be green with envy, red with anger, or white with fear. These emotionally charged uses for colors would seem to indicate that there might be a psychological connection between color and emotions, as revealed by Valdez and Mehrabian (1994). Other paths than emotion may exist. Mood is one of these. It's an affective state of mind less likely to reach our conscience which lasts longer than emotions but is less intense (Forgeas, 1999). Age can also moderate the perception of color and any culture has its own perception of color: For example, people in northern countries wear white clothes at their weddings while Asian wear black ones.

Color is a design element of every modern computer interface, including E-commerce sites. E-commerce sites are frequently updated, changing many design elements, including media, navigation, and colors (Benbunan-Fich and Altschuller 2005). These updates represent a considerable investment in time, effort, and money, with the goal of improving the user experience and thereby increasing sales. In practice, website updates are designed ad hoc, like many IT projects, and often fail because of a lack of understanding of what the user wants (The Standish Group 1995). Such updates often meet user resistance, requiring retreat to an older version. These update rejections waste considerable development money, but more importantly can cause the company to lose customers.

E-commerce sites need to elicit the best possible first impression, as users form a rapid impression of a website, which tends not to change (Tractinsky et al. 2006). One of the first things that users notice about a site is its general color scheme. Since people have an affective reaction to color, it seems reasonable to expect that much of a user's initial affective reaction to a website is due to its color. If we know what colors generate the most positive first impression, we can improve users' overall attitude toward the site.

The lack of understanding about the best ways to implement a website which result in their ad hoc updates is due to gaps in the scientific research. While many studies have focused on how reactions to an e-commerce site can change use or purchasing intentions (e.g (Gefen et al. 2003; Pavlou and Gefen 2004), fewer have examined how site design can create affective reactions. When color is included as a design element, it is considered in terms of cultural prevalence (Lo and Gong 2005), readability (Kubilus 2000), the presence or absence of color of any type (Benbasat and Dexter 1986), or as general heuristics (Wang and Emurian 2005), without examining the full gamut of color and why it matters. One of the few studies to consider broader types of color focuses on how saturation and brightness can increase recall (Pelet 2010). At the same time, much psychology and marketing literature examines color's effect upon mood or affect (Brengman and Geuens 2004). This literature suggests that color should be an important determinant of website interaction. However, outside of a few works (e.g., Gorn et al. 1997, Gorn et al. 2004, Pelet 2010), little research has examined such a link, and almost none has studied the psychological mediators which translate physical color into attitudes about a website.

We suggest several mechanisms by which color attributes affect website perception and use. While there are likely paths due to cognitive aspects, and related issues such as readability and color sensitivity, for this study we have chosen to focus on possible emotional paths. Color components seem to activate elements of the human arousal system, which may then affect our trust in an e-commerce site. Arousal will tend to increase the notice of details, which, depending upon the type of arousal, can increase trust (energetic arousal) or reduce trust (tension arousal). This result is similar to Mandel and Johnson (2002). In that research, appearance modifies choice depending upon the color of the screen background used to display attractive products- using an orange red with a car (exciting colors), green with dollars (relaxing colors), and blue with white clouds for a sofa. These two proximal effects of color result in distal effects on perceived quality of an e-commerce site. Because these different types of arousal work in opposition, the question of optimal color choice is quite complex.

Our research question is: Can color affect the perception of quality of an e-commerce website, and are emotional arousal and trust the mechanisms of this effect? This paper is structured as follows. We examine the background literature in more detail, and define some terms. We develop a new theory and model to help fill the gaps that we have identified. We present an exploratory study to validate our method and measures. We present and discuss the results, and examine the implications for practice and research.

Related Work

In order to develop the background of this research, we begin with a definition of color and then begin to integrate the psychology and marketing literature stream relating color and affect, with the information systems literature on website design.

Before we begin to talk at all about "color", we must understand how to talk about it. Color is a psychological construct based upon the visual response of the neural network attached to the light receptors in the retina. The color ascribed to this visual perception depends on the physical wavelength of the light reflected (or generated) by the object, the physiology of the recipient, and the individual's learned color response (Munsell 1905). To try to cope with this complexity, a number of methods of color naming and mapping have been created. The most common is the Munsell model, which maps colors onto a three dimensional cylinder of hue, chroma or saturation, and brightness (Munsell 1905). Each of these components can affect an individual in different ways.

Munsell (1905) defines the components of color that follow throughout this paragraph. Hue is the actual tint, e.g. red, yellow, or green-- the wavelength of the light that is perceived by the eye. Saturation is purity of color-- the extent to which the color is composed of a single wavelength of light. A bold red has higher red saturation than pink does. Brightness is the overall intensity of the color. White has low saturation (all wavelengths present) but high brightness. Black has a brightness of 0, making its saturation unknowable. Laser light, which is light of one wavelength, is completely saturated (and generally very bright). Our experience of these colors happens at the intersection of the physical perception of light in the blue-violet, green, and yellow-green spectrum, the neural network that processes that information, and our learned color definitions.

There are several theories which describe how we perceive colors based on those inputs. For this research, we use the opponent theory perspective, which holds that colors are sensed as the difference between the activation levels of the sensing neurons. In this system, yellow is opposed to blue, red is opposed to green, and black is opposed to white. We choose this perspective because the affective reactions to oppositions of these pairs of colors seems to be much stronger than the affective reactions described in other theories. This suggests that opponent color theory is more suited to evaluate emotional responses to color (Pokorny and Smith 1986).

Our emotional reactions to colors seem to be linked to all three color attributes (Guilford & Smith, 1959). Some links are rooted in culture; some values hold across cultures. Universal reactions include the following. Blues are fairly universally seen as relaxing, pleasant, calm, and reassuring. Yellow, which is in opposition to blue, is exciting, unpleasant, active, and causes wariness (Adams and Osgood 1973). Both blue and yellow are linked to elements of arousal which control excitement. Saturation is also linked to excitement, with higher saturation leading to higher excitement. Brightness has an inverse link to tension or anxiety; the darker something is, the more tense or anxious the viewer gets (Walters et al. 1982).

Hue has been examined most of the three components-- especially red and blue. Respondents were likelier to buy products, and buy them more quickly, in blue environments rather than red (Bellizzi and Hite 1992). When products are presented on a blue background, they are evaluated more favorably than when presented on a red background (Lichtlé 2007). Red generates avoidance behavior, leading to a reduction in achievement (Elliot et al. 2009). In general, people prefer shades of blue or green to shades of red or yellow (Guilford 1934; Guilford and P. C. Smith 1959). Blue hues induce more relaxation than red or yellow hues, leading to a perception of increased website speed (Gorn et al. 2004).

Brightness and saturation have been less thoroughly studied; however, the existing results seem stronger than the results of hue studies. In general, people have a more positive affective reaction toward high brightness and high saturation (Guilford and Smith 1959) When measured on the pleasure/arousal/dominance (PAD) scale (Mehrabian and Russell 1974), saturation increased all three components, while brightness increased pleasure but decreased arousal and dominance (Valdez and Mehrabian 1994). Increased brightness and saturation increased the favorable ratings of a product (Middlestandt 1990). Higher saturation in advertisements increased excitement; higher brightness increased relaxation. Both effects increased the liking for the advertisement (Gorn et al. 1997; Lichtlé 2007). Higher saturation and lower brightness increased relaxation, leading to a perception of faster website download (Gorn et al. 2004). Finally, higher saturation and brightness improved recall of site information and intention to buy (Pelet 2010).

Gorn et al. (1997) and Gorn et al. (2004) examined how colors affected consumers' attitudes through the phenomenon of arousal, using the Thayer (1986) scale. This scale measures arousal with two factors representing

energetic arousal, or excitement, and anxious arousal, or tension. In this way of defining arousal, the ranges of the two dimensions are excitement to boredom for energetic arousal, and tension to relaxation for anxious arousal (Thayer 1986). Although other scales exist, this definition of arousal seems more tied into emotional than cognitive appraisals, and is therefore the one we use in this research.

Previous attempts have been made to integrate color into website use the hue or saturation as a direct antecedent to trust. The primary examinations of the relationship of color to website use distinguished between warm (red, yellow) or cool (blue) colors and pastels (low saturation) versus bolds (high saturation) (Wang and Emurian 2005; Yang et al. 2005). However, a human's perception of color is more complex than warm or cold, pastel or bright. Early research into colors showed that there is a spectrum of reactions to hues, saturation levels, and brightness levels, forming a complex set of curves (Guilford and P. C. Smith 1959). It is likely that these complex surfaces have a more complex relationship with trust. Pelet and Papadopoulou (2010), in an exploratory qualitative study on how the colors of an e-commerce website can affect trust, found color to be an important e-commerce interface factor influencing customer trust and the specific trust beliefs in the benevolence, competence, integrity and predictability of an online vendor. According to their findings, colors which are low in saturation and brightness infuse trust in an online vendor and customer beliefs in the vendor's benevolence, competence, integrity and predictability, whereas colors which are vivid and high in saturation have a negative effect on trust and its four components

While color has been considered as an antecedent to trust,t seems likely that some psychological construct activated by color mediates the change in trust. Since studies have shown that colors increase arousal, we expect that this arousal could cause the observed changes in trust.

Trust has been shown to be critical for the success of an e-commerce website (McKnight et al. 2002a, 2002b). Previous studies have shown that some website design aspects have been shown to increase trust; e.g. the existence of user feedback or branding can increase website acceptance, depending on cultural values (Sia et al. 2009). Finally, it has been shown frequently that increased trust in an e-commerce site leads to greater intention to use the site (McKnight et al. 2002b). Thus, we plan to show that changes in the brightness and/or saturation of an e-commerce website's color will have direct effects on the success of that website. It should be noted that McKnight et al. (2002a, 2002b), Gefen et al. (2003), and Gefen and Straub (2004) have modeled trust as a second-order construct formed from the four components of benevolence, competence, integrity, and reliability. We will use trust in that second-order sense in our theory, but note that the measurement and analysis will be based upon trust as a second-order factor composed of those four elements.

Theoretical Model

We use Munsell's (1905) definitions to describe the colors used on a website. Each color is composed of three components: hue, saturation, and brightness. Each component affects dimensions of arousal in an individual (Brengman and Geuens 2004). We will not attempt to study the effects of varying hues in the present work, preferring to study the less familiar components of saturation and brightness.

Our first three hypotheses concern potential changes in arousal levels resulting from various color component variations.

In studies of color and arousal in traditional media, saturation has been shown to increase the energetic component of arousal (the excitement-boredom axis). Higher saturation colors lead to higher affective reactions (greater excitement). Since saturation is a description of how much of the "pure" (single wavelength) color is perceived by the observer, higher saturation will tend to polarize reactions more strongly, since there will be a strong primary, and weak opponent, signal for the color internal to the brain. This clearer and stronger signal should be more noticeable and more likely to provoke a reaction. The effect should be that higher saturation enhances the affective reaction to the color. When the affective reaction is an increase in energy, a more saturated color should increase the energy more. When a color would normally cause tension, a more highly saturated color will cause more tension. Pastel colors (low saturation) should cause less energetic arousal than bold colors (high saturation).

H1A: An increase in color saturation will increase energetic arousal

H1B: An increase in color saturation will increase tension arousal

In conventional media brightness has been shown to affect tension. Since a brighter color is easier to see, we more quickly determine its hue, and come to a conclusion about its threat or opportunity level. When a color is dimmer,

we strain more to determine its hue, and that determination may take longer. That period of uncertainty about hue causes anxiety. The dimmer a color it is, the more anxious it makes us. However, this perceptual difference is not expected to change the level of energetic arousal a color causes, and thus brightness should only affect tension arousal.

H1C: An increase in brightness will increase tension arousal

Our next two hypotheses concern potential changes in trust levels resulting from changed arousal levels.

Energetic arousal concerns feelings of excitement or eagerness. The opposite pole is boredom. When we get excited, we tend to be happier and look more favorably upon our surroundings. Effectively, we become more optimistic. We are likelier to think well of those around us, and more positively about the things we are interacting with. Thus, in an e-commerce site, an increase in excitement should lead to an increase in positive perceptions about the site. One of the major perceptions we have about a site is its trustworthiness. Thus, we expect that as energetic arousal increases, trust in the e-commerce site will increase.

H2A: An increase in energetic arousal will increase trust in an e-commerce site.

An increase in tension arousal increases the attention that one pays to the environment. When you become tense, you become more wary of your surroundings; tension implies that there is a threat. Tension makes you more alert and more likely to make quick judgments about the situation. You want to avoid being surprised; if you can predict accurately what will happen in the situation, you can react more quickly to alleviate the threat. It is generally safer to consider something a threat, and be proved wrong, than to assume that something is safe and be incorrect. You become warier. This wariness makes you less likely to trust the motives of those around you, since trusting someone, by definition, means that you put yourself at risk of betrayal by that person. In this case, you are interacting with an e-commerce website. The higher your tension arousal, the warier you are. When you are more wary, you are less willing to put yourself at risk. Thus, you would be less likely to trust the website when your tension arousal is high.

Thus, we expect that website colors that lead to higher tension arousal will decrease trust in the website.

H2B: Increased tension arousal will lead to reduced of trust in the e-commerce website

Our final hypothesis concerns the effect of changes in trust levels. That increased trust increases the perceived quality of a site is well known, and we include this in our model to validate it in the context of the known nomological net.

H3: An increase in trust will increase the perceived quality of the e-commerce site.

Figure 1 below summarizes our model.

Method

To evaluate our hypotheses, we designed a controlled randomized experiment. This exploratory laboratory study was done at large university in the southern United States, using student volunteers as respondents, and employing a mock e-commerce website. Students were recruited from the management and information technology departments and offered extra credit for their participation. They were told that they would be participating in a study on how website design elements affect user reactions. Color was not mentioned. The experiment was set up with 16 conditions, summarized in Table 1. The design is a 2x2x2x2 between subjects experiment with random assignment; the treatments were high and low saturation foreground, high and low saturation background, high and low brightness foreground, and high and low brightness background. "High" and "low" are relative to each other. In fact, the difference between the levels was kept fairly small to ensure that there were no differences in readability between the conditions. The RGB values for the foreground were 0x006500, 0x009100, 0x1e481e, and 0x2b662b. The RGB values for the background were 0xffffd6, 0xffffac, 0xf3f3e2, and 0xe7e7c5. These correspond to a fairly dark green foreground and a fairly light yellow background. Hill and Scharff (1997) showed that the sharp contrasts of this combination provided users the fastest reading speed possible among the tested hues. The foreground color was used for the site text for descriptions, names, links, and all textual information. The site background was the only other color present except for the images of the CD covers. The catalog of available CDs was identical for all respondents. Exactly which CDs they viewed depended upon their own interest.

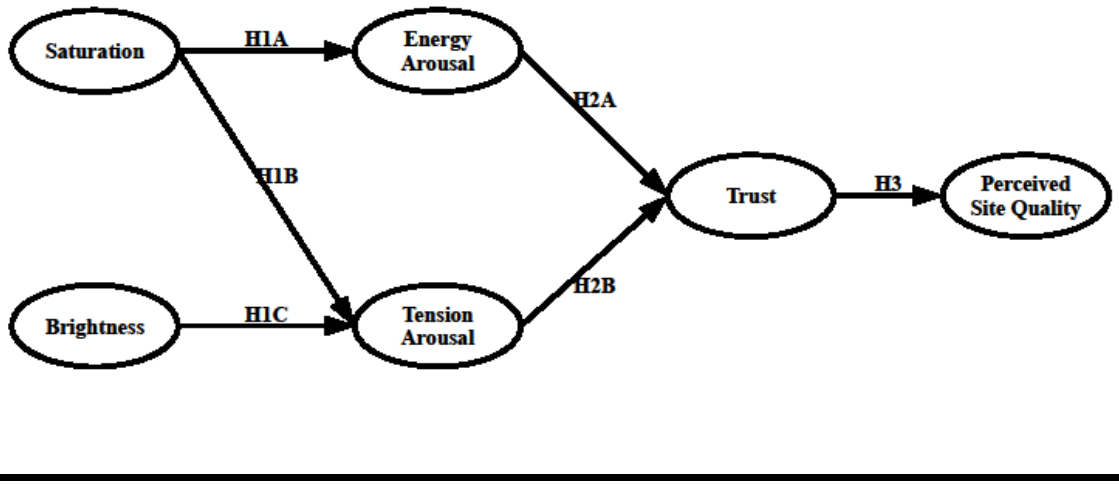

Figure 1: Research Model

Since the CD covers varied in their color, it is a natural concern that these covers might confound our results. However, this is actually no different from any other individual difference. Each respondent was free to view CDs of their choosing, which may have varied in color from the CDs viewed by other respondents. However, all respondents chose the CDs they viewed from the same pool of 57 CDs. Since they were assigned to treatment conditions randomly, their possible predilection to choose CDs of any particular color will have been spread randomly across the treatment groups. Any bias due to overall coloring, or attractiveness or popularity of any particular CD or category, would serve to reduce the differences between the treatment groups. To see this, consider if, instead of CD covers, the respondents merely chose a color for a square in the middle of their screen from a set of 57 possible colors. While any particular individual might choose a color that might reinforce their randomly chosen treatment, they might also have chosen a color that reduces the effect of the chosen treatment. At worst, this ability to see another color will increase the standard error in our measures, reducing the power of the study. This is a statistical conclusion validity threat in that it increases the chance of a Type II error. The chance of a Type I error is unchanged. In practical terms, this means that our study should be conservative in terms of finding effects due to color, in that any effects that we find are fairly sure to exist; a failure to find an effect may be due to Type II error.

On arrival at the study site, participants were instructed to browse the study website and then fill out a survey. The study website is a locally constructed mock e-commerce site which purports to sell CDs. Hues were fixed with a green foreground and yellow background, and brightness and saturation ranges were limited to avoid a confound due to readability (Itten and Birren 1970). Respondents browsed the website for a period of time of their choice (typically 10-20 minutes). After they examined two individual items from the catalog, a link appeared allowing them to respond to the survey. They could follow the survey link at any subsequent point. An example screenshot from the site is in Figure 2. The background (light yellow) and foreground (dark green) are the elements which are manipulated in the experiment. The CD cover comes from a set of 57 available. The same 57 CD covers are available to all users, though they will see only the ones they choose. The other components of the page (e.g. musicashop.net logo) are identical in all conditions.

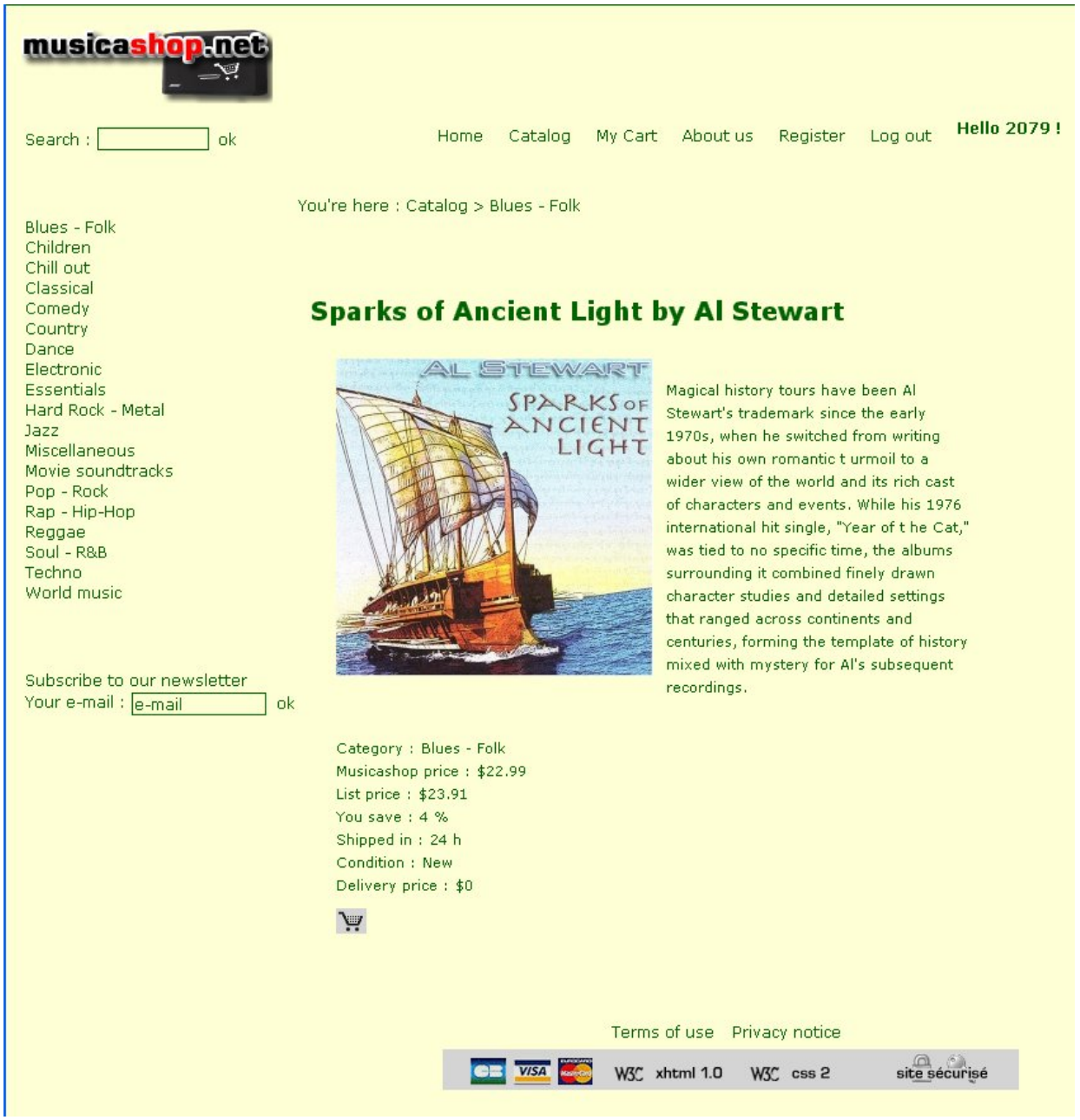

Figure 2: Example screenshot from site

The study was conducted in a room in which the ambient lighting was carefully controlled, set to 1000 lux with a temperature of 4500 Kelvin. The LCD monitors were calibrated and ICC profiles set to faithfully and uniformly present the colors, at the recommended settings for an LCD monitor: 6500 Kelvin temperature, 120 luminance, and 2.2 gamma. Both ambient light and monitor calibration were performed with an X-Rite i1 device using Eye-One Match software version 3.6.1. LCD monitors do suffer from degradation of image when not viewed from the correct angle. Hollands et al. (2002) conducted experiments in order to determine the best visual search performance for tactical symbols due to perceptual differences in displays. They noted differences between liquid-crystal (LCD) and cathode-ray tube (CRT) displays. They showed that participants in their experiment who had to search for navy tactical display symbols on a map background were less able to detect colored targets viewed off-axis on a LCD

screen. To deal with this, we positioned the monitor and chair so that the respondent was within the acceptable range for most effective viewing. However, it is not realistically possible to ensure that respondents did not shift in their chair, adjust the monitor, or other such changes which might shift the respondent from the optimal viewing setup. However, we expect that people will usually make such adjustments to improve their view of the monitor; and, at any rate, random assignment should reduce the influence of any effect resulting from this.

The saturation and brightness of the foreground and background colors were assigned to respondents randomly. Brightness and saturation were manipulated independently, resulting in 16 possible conditions, summarized in Table 1.

Table 1: Experiment Conditions

Foreground Saturation	High								Low							
Foreground Brightness	High				Low				High				Low			
Background Saturation	H		L		H		L		H		L		H		L	
Background Brightness	H	L	H	L	H	L	H	L	H	L	H	L	H	L	H	L
Condition	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

We used previously validated scales for all our measurements. We measured the four components of trust (benevolence, competence, integrity, and predictability) using items adapted from McKnight et al. (2002a) and Gefen and Straub (2004). To measure energetic and tension arousal, we used the Thayer (1986) scale. We used the propensity to trust and perceived site quality scales developed by McKnight et al. (2002b). Table 2 lists the measures used, and provides the items.

Table 2: Measures

Energetic Arousal	**Thayer 1986**
respondents were asked "Mark the extent to which you feel", selecting from "definitely feel", "feel slightly", "cannot decide", "definitely do not feel", for each of the following adjectives: Energetic, Lively, Active, Vigorous, Full of Pep, Sleepy, Drowsy, Tired, Wide-Awake, Wakeful (energy00-energy09)	
Tension Arousal	**Thayer 1986**
respondents were asked "Mark the extent to which you feel", selecting from "definitely feel", "feel slightly", "cannot decide", "definitely do not feel", for each of the following adjectives: Tense, Clutched-Up, Fearful, Jittery, Intense, Still, At Rest, Calm, Quiet, Placid (tense00-tense09)	
Trust: Benevolence	**Adapted from Gefen and Straub 2004 and McKnight *et al.* 2000a**

Table 2: Measures	
benevolence00: I believe that MusicaShop.net intentions are benevolent benevolence01: I believe that MusicaShop.net is well meaning benevolence02: I believe that MusicaShop.net would act in my best interest. benevolence03: If I required help, MusicaShop.net would do its best to help me. benevolence04: MusicaShop.net is interested in my well-being, not just its own. benevolence05: MusicaShop.net would not knowingly do anything against my interest benevolence06: MusicaShop.net would not take advantage of me	
Trust: Competence	**Adapted from Gefen and Straub 2004 and McKnight *et al.* 2000a**
competence00: MusicaShop.net knows about music competence01: MusicaShop.net is capable of providing excellent service competence02: MusicaShop.net is competent and effective in selling music. competence03: MusicaShop.net performs its role of selling music very well. competence04: Overall, MusicaShop.net is a capable and proficient music seller. competence05: In my opinion, MusicaShop.net is capable of meeting its obligations as an online music seller competence06: MusicaShop.net is capable of fulfilling its promises (for example, regarding the delivery of orders within the time promised)	
Trust: Integrity	**Adapted from Gefen and Straub 2004 and McKnight *et al.* 2000a**
integrity00: Promises made by MusicaShop.net are likely to be reliable integrity01: MusicaShop.net is truthful in its dealings with me. integrity02: I would characterize MusicaShop.net as honest. integrity03: MusicaShop.net would keep its commitments. integrity04: MusicaShop.net is sincere and genuine. integrity05: MusicaShop.net does not make false claims integrity06: MusicaShop.net is not always honest with me	
Trust: Predictability	**Adapted from Gefen and Straub 2004 and McKnight *et al.* 2000a**
pred00: I am quite certain about what MusicaShop.net will do pred01: I am quite certain what to expect from MusicaShop.net pred02: MusicaShop.net behaves in a consistent manner pred03: MusicaShop.net behaves as expected pred04: MusicaShop.net does the same thing every time the situation is the same pred05: I seldom know what MusicaShop.net will do in a given situation	

Table 2: Measures	
Propensity to Trust	**McKnight *et al.* 2002a**
I usually trust people until they give me a reason not to trust them. I generally give people the benefit of the doubt when I first meet them. My typical approach is to trust new acquaintances until they prove I should not trust them.	
Perceived Site Quality	**Adapted from McKnight *et al.* 2002b**
qual00: Overall, this site worked very well technically. qual01: Visually, this site resembled other sites I think highly of. qual02: This site was simple to navigate. qual03: On this site, it was easy to find the information I wanted. qual04: This site clearly showed how I can contact or communicate with MusicaShop.net.	

Once the respondents answered the survey questions, they were given an Ishihara pseudoisochromatic test (Waggoner 2005) to determine if they had any defect in their color vision. We performed the Ishihara test at the completion of the survey in order to avoid hypothesis guessing by the respondents. We deleted data for three respondents with color vision defects. There were no respondents with other color defects. Additionally, because the survey could not require responses to all questions, we had to discard 42 responses in which fewer than half of the survey questions were answered. The remaining respondents answered virtually all the survey questions, and we judged them to be complete enough to use in the analysis. At the end, we had 109 usable responses total. Because this was not a large enough sample to examine the full factorial design, we tested for interaction effects between the elements of saturation and brightness. The interaction effects were not significant, so we pooled data for each of the conditions, resulting in four independent variables for use in the PLS analysis, with roughly 55 data points in each category.

Results

Our results show good statistical validity. Analysis of the data was performed using SmartPLS version 2.0M3 (Ringle et al. 2005). Since there were still a few missing cells in our final 109 data points, we used the mean replacement option of SmartPLS to fill the cells for analysis. The first run indicated several items, particularly in the Thayer scale, which did not load well on their constructs. As this is an exploratory examination of the model, we removed those items which loaded poorly (less than 0.7 on their intended construct). We eliminated items quality01 and quality04 from the Quality scale. The Thayer scales have not received the same kind of thorough statistical grounding as more recent scales. It is thus not surprising that it might have more issues. As a result, we retained only the following items for the Thayer scales: energy05-energy08 and tense01-tense03. The resulting factor loadings for items on their construct were satisfactory. All items except for benevolence06 had loadings over 0.7. Benevolence06 had a loading of 0.65; however, the reliability of the measure was still good, and the measure is established, so we retained that item. Crossloadings were generally below 0.4. The main exception is that there was some crossloading in the trust component measures. However, the convergent and discriminant analysis for these measures was satisfactory, and they are well accepted, so we did not attempt to pare items. (For reasons of space we do not include the first-order measurement assessment of the trust measure; this information is available from the lead author.) The resulting AVE, composite reliability, and Cronbach's α for each factor is presented in Table 3, AVE and Reliability. Convergent validity is indicated when AVE > 0.5 and reliability > 0.7 (Fornell and Larcker 1981). We report both Cronbach's α and composite reliability, because while Cronbach's α gives a more conservative measure of reliability than composite reliability, composite reliability is considered to be more accurate in SEM analysis (Raykov 1998). In this case, the Cronbach's α and the composite reliability for all measures exceeds the 0.7 standard. All measures also have AVE >0.5. Together these indicate good convergent validity. Because of our use of Trust as a second-

order formative factor, it is not possible to provide reliability figures for this measure.

Table 3: AVE and Reliability

Measure	AVE	Composite Reliability	Cronbach's α
Tension Arousal	0.62	0.83	0.70
Energetic Arousal	0.83	0.95	0.93
Quality	0.66	0.85	0.74

Table 4, Correlations, presents the correlations between the latent variables. The diagonal in this table contains the square root of the AVE. The Fornell-Larcker test for discriminant validity consists of comparing the square root of the AVE to the correlations between the latent variables. If the correlations between a construct and all other constructs is less than the square root of the AVE for that construct, we demonstrate discriminant validity (Fornell and Larcker 1981). All measures meet this test, and we conclude that we have discriminant validity for these measures.

Table 4: Correlations (Diagonal is Square Root of the AVE)

	Energetic Arousal	Quality	Tension Arousal	Trust
Energetic Arousal	0.91			
Quality	0.13	0.81		
Tension Arousal	-0.23	-0.24	0.79	
Trust	0.18	0.65	-0.30	1

Figure 3 and Table 5 present the PLS results. The paths are marked with standardized path coefficients, their t-values, and their p-values. The t values were calculated from a bootstrap using 109 samples (the number of data points) and 1000 iterations, resulting in 108 degrees of freedom. The latent variable labels contain the R^2 explained by the model.

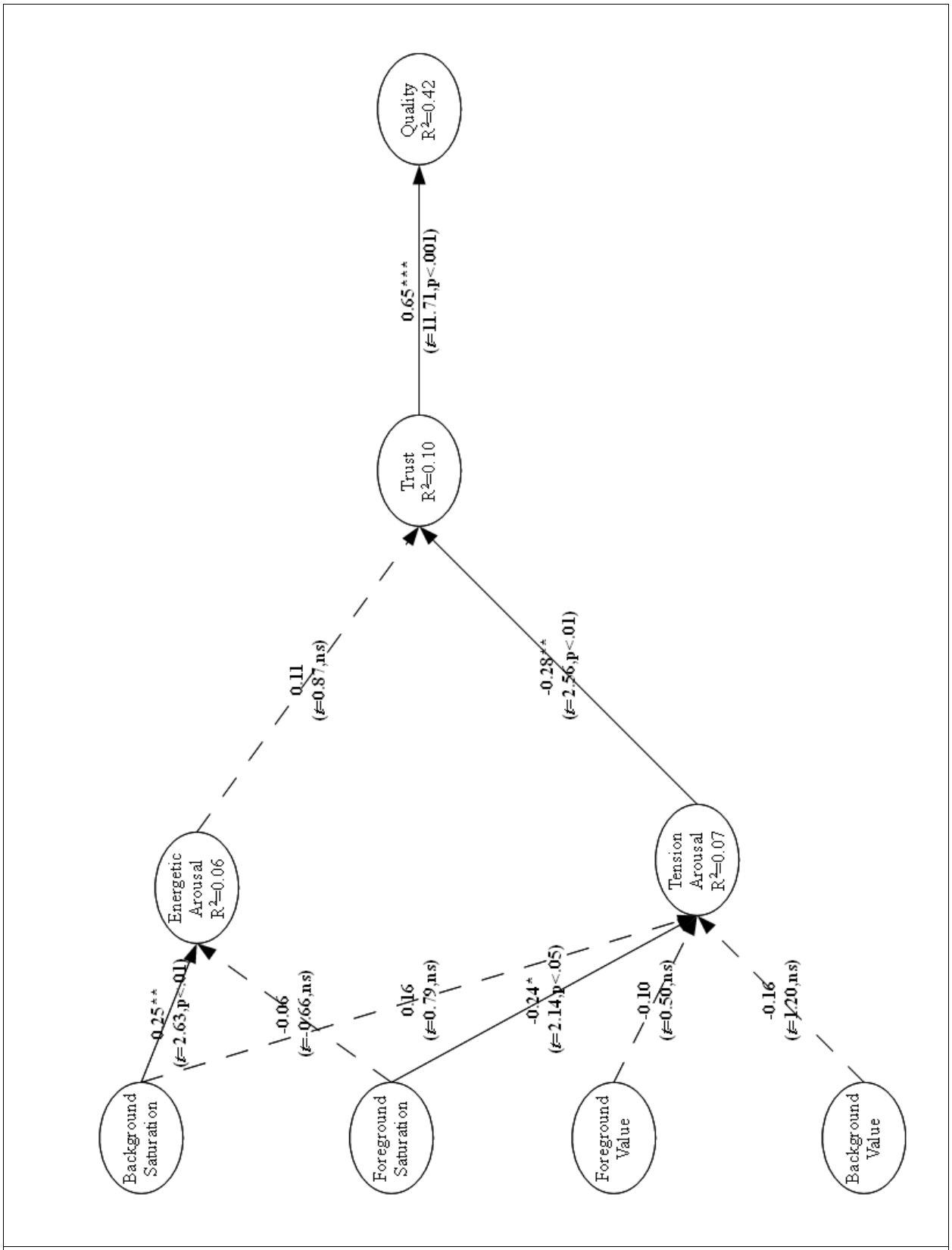

Figure 3: PLS Results

Table 5: PLS Results

Hypothesis	Independent Variables	Dependent Variables			
		Energetic Arousal	Tension Arousal	Trust	Perceived Quality
H1A (energetic) and H1B (tension)	Background Saturation	0.25**	0.16		
	Foreground Saturation	-0.06	-0.24*		
H1C	Background Brightness		-0.16		
	Foreground Brightness		-0.10		
H2A	Energetic Arousal			0.11	
H2B	Tension Arousal			-0.28**	
H3	Trust				0.65***
	Dependent Variable R^2	0.06	0.07	0.14	0.42

(+p<0.10 *p<0.05 **p<0.01 ***p<0.001)

We turn now to a narrative of our results. H1A hypothesized that increased saturation would increase energetic arousal. This hypothesis held true for background saturation, but not foreground; H1A is partially supported. H1B reverses this outcome; foreground saturation is significant for tension arousal, but not background saturation; H1B is partially supported. The increased attention paid to the foreground rather than the background color may explain this difference. Tension arousal is likelier when there is a specific stimulus; thus, foreground color (i.e. text), which is more salient to an e-commerce site user, is a more immediate "threat" than background color. Excitement tends to be more broad-based, and so might be more affected by background color, which is not the direct focus of the viewer. It is also possible that the differences in saturation were not large enough to provoke a measurable difference, or that the sample size was inadequate.

H1C held that increased brightness would lead to decreased tension arousal. Brightness did not significantly increase tension arousal, so this hypothesis was not supported.

H2A held that an increase in energetic arousal would increase trust. We did not find a significant relationship here. Upon examining the results in more detail, PLS loaded the trust measure primarily from the benevolence component. Energetic arousal was not significantly related to benevolence, but was significantly related to the other components; thus we consider H2A to be partially supported. H2B held that an increase in tension arousal would decrease trust. This path is significant, so hypothesis H2B is supported.

The final hypothesis, H3, was that increased trust would lead to an increase in perceived quality. This path was significant, with a very high path coefficient, so H3 is supported.

While the low R^2 and path coefficients might be purely from lack of statistical power, it is also possible that it is due to mediational paths that are not included in this model. The small sample size does not permit running a full ANOVA to test this. However, one-way ANOVAs of the foreground and background conditions considered separately were not significant. This suggests that, while there may well be alternate paths which are important, the main issue causing our results is lack of statistical power, or insufficient manipulation of the independent variable.

Limitations

Our exploratory study sample, like many of its type, was drawn from a university student population, which is likely to be more homogeneous than the general population. Because it is a convenience sample, it is likely that we have not yet sampled enough members of different cultures, who would have potentially differing attitudes toward colors. The sample size was also fairly small, which may have resulted in Type II errors when evaluating our hypotheses.

As has been noted in (Stone-Romero and Rosopa 2004; 2008) tests of mediation, which is implicitly tested in PLS analysis, are subject to a number of analytical and experimental issues. In particular, while a true experiment in a special purpose setting can provide strong evidence for mediation (Stone-Romero and Rosopa 2004), a test from a single experiment is weaker than a test by two experiments. This is because it is not possible to state analytically from the results of one experiment whether the mediator causes the dependent variable, or vice versa. However, additional studies which manipulate the two intermediary mediators would certainly strengthen our results.

Only a small portion of the extremely large color space can be explored in any study. Further studies to confirm the effects across other parts of the color space will help to support the validity of this theory.

We did not manipulate the hue in this experiment; our sample size was not large enough to allow for another major variable. However, the prior literature leads us to believe that hue will be important component of this model.

In order to ensure that readability issues did not confound our results, we limited the saturation and brightness to a relatively narrow range, which reduces their effect size. A better approach might be to measure readability, and allow the color characteristics to vary more widely.

At this point, only exploratory study results have been obtained. The data are sufficient to indicate the possibility of the model being supported, and to support the validity of the measures themselves. Because of the small sample size, the statistical power is low, and we may have Type II errors (failure to find an effect when, in fact, there is an effect). Testing of the full factorial model will require a larger sample size.

Contributions and Future Work

The present work demonstrates the value of integrating the extensive literature streams of color's effect on psychological affect with the information systems literature examining design characteristics that create trust in (and thus use of) a website. Our research also begins to explain one of the psychological mechanisms by which color affects trust in an e-commerce website. It shows that components of colors may create tension and energetic arousal, and that tension and energetic arousal may affect attitudes toward the e-commerce site.

Higher tension arousal leads to both better recall and lower trust in an e-commerce site. Better recall leads to increased likelihood of use of the website (Pelet 2010), but reduced trust decreases the likelihood of use of the website (McKnight et al. 2002b). A better understanding of the interaction between these two opposite effects of tension arousal would be interesting.

One avenue for research concerns the effect of different hues. A single viewer may react differently to different hues; two or more viewers from different cultures might react differently to the same hue, since cultural interpretations of hues vary and might moderate the effects seen. Further, this study examined only a small variation in saturation and brightness. Larger differences will likely have a larger effect, and may prove of value in website design.

It could be argued that increased brightness might be expected to increase energetic arousal as well as decreased tense arousal. Exploring this would improve our understanding of how these forms of arousal can be created by the various components of colors. It is also possible that there may be interaction effects from the various components of color. For instance, high saturation in the foreground and the background might lead to a greater effect than the simple addition of the two. It is also possible that saturation and brightness might interact with each other, or with

hue. The color space is rich and varied, and because a single study can only explore a small part of the space, much work will be needed to provide truly generalizable results.

Questions of readability have plagued the on-line world from the beginning. Since color is an important part of readability, incorporating the existing literature on how readability affects consumer behavior would improve this theory.

Finally, an age-, gender-, and culture-neutral color space against which reactions could be measured would be a valuable contribution to further research on color and website use.

Conclusion

In this work, we have begun the exploration of the relationship between psychological and marketing color theory, and information systems theories on trust. We have done an exploratory study which confirms our hypotheses about the relationships between color components and arousal states, and between arousal states and trust. We are able to provide a preliminary guideline for color design of sites which should help to ensure that newly designed sites meet their objectives. In particular, the present research suggests that e-commerce sites with highly saturated, brighter colors will be perceived by users to be higher quality than paler, dimmer sites. First impressions still count, and improved color design in e-commerce sites has tremendous potential to increase user trust and thereby reduce the waste from rejected sites and site upgrades. As more and more companies continue to transition into on-line environments, further exploration of the color domain for websites and other virtual spaces can help ensure that their customers will respond positively to the companies' online presence.

References

Adams, F., and Osgood, C. 1973. “Cross-Cultural Study of Affective Meanings of Color,” Journal of Cross-Cultural Psychology (4:2), pp. 135-156.

Bellizzi, J. A., and Hite, R. E. 1992. “Environmental Color, Consumer Feelings, and Purchase Likelihood,” Psychology and Marketing (9:5), .

Benbasat, I., and Dexter, A. S. 1986. “An Investigation of the Effectiveness of Color and Graphical Information Presentation Under Varying Time Constraints.,” MIS Quarterly (10:1), March, pp. 59-83.

Benbunan-Fich, R., and Altschuller, S. 2005. “Web Presence Transformations in the 1990s: An Analysis of Press Releases,” IEEE Transactions on Professional Communication (48:2), pp. 131-146.

Brengman, M., and Geuens, M. 2004. “The Four Dimensional Impact of Color on Shopper's Emotions.,” Advances in Consumer Research (31:1), January, pp. 122-128.

Elliot, A. J., Maier, M. A., Binser, M. J., Friedman, R., and Pekrun, R. 2009. “The Effect of Red on Avoidance Behavior in Achievement Contexts,” Personality and Social Psychology Bulletin (35:3), pp. 365.

Forgeas J P (1999) Network theories and beyond, éds. T. Dalgleish and M. J. Power, Handbook of Cognition and Emotion, 591-612, Chichester, Wiley.

Fornell, C., and Larcker, D. F. 1981. “Evaluating Structural Equation Models with Unobservable Variables and Measurement Error.,” Journal of Marketing Research (18:1), February, pp. 39-50.

Gefen, D., Karahanna, E., and Straub, D. W. 2003. “Trust and Tam in Online Shopping: An Integrated Model,” MIS quarterly (27:1), pp. 51-90.

Gefen, D., and Straub, D. W. 2004. "Consumer Trust in B2C E-Commerce and the Importance of Social Presence: Experiments in e-Products and e-Services," Omega (32), pp. 407-424.

Gorn, G. J., Chattopadhyay, A., Sengupta, J., and Tripathi, S. 2004. “Waiting for the Web: How Screen Color Affects Time Perception.,” Journal of Marketing Research (41:2), May, pp. 215-225.

Gorn, G. J., Chattopadhyay, A., Yi, T., and Dahl, D. W. 1997. “Effects of Color as an Executional Cue in Advertising: They're in the Shade.,” Management Science (43:10), October, pp. 1387-1400.

Guilford, J. P. 1934. “The Affective Value of Color as a Function of Hue, Tint, and Chroma.,” Journal of Experimental Psychology (17:3), June, pp. 342-370.

Guilford, J. P., and Smith, P. C. 1959. “A System of Color-Preferences,” The American Journal of Psychology

(72:4), December, pp. 487-502.
Hill A. et Scharff L. V. (1997), Readability of websites with various foreground/background color combinations, font types and word styles, Proceedings of 11th National Conference in Undergraduate Research, 2, 742-746.
Hollands J. G., Parker H. A., McFadden S., Boothby D. R. (2002), LCD versus CRT displays: A comparison of visual search performance for colored symbols, Human Factors and Ergonomics Society, vol. 44, no2, pp. 210-221.
Itten, J., and Birren, F. 1970. The Elements of Color, John Wiley & Sons.
Kubilus, N. J. 2000. "Designing an E-Commerce Site for Users," Crossroads (7:1), pp. 23-26.
Lichtlé, 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), February, pp. 37-62.
Lo, B. W. N., and Gong, P. 2005. "Cultural Impact on the Design of E-Commerce Websites: Part I–Site Format and Layout," Issues in Information Systems (6:2), pp. 182-189.
Mandel, N., and Johnson, E. J. 2002. "When web pages influence choice: Effects of visual primes on experts and novices," Journal of Consumer Research. 29(September), pp.235-245
McKnight, D., Choudhury, V., and Kacmar, C. 2002a. "Developing and Validating Trust Measures for E-Commerce: An Integrative Typology," Information Systems Research (13:3), September, pp. 334-359.
McKnight, D., Choudhury, V., and Kacmar, C. 2002b. "The Impact of Initial Consumer Trust on Intentions to Transact with a Web Site: A Trust Building Model," Journal of Strategic Information Systems (11:3-4), December, pp. 297-323.
Mehrabian, A., and Russell, J. A. 1974. An Approach to Environmental Psychology, MIT Press.
Middlestandt, S. E. 1990. "The Effect of Background and Ambient Color on Product Attitudes and Beliefs," Advances in Consumer Research (17:1), January, pp. 244-249.
Munsell, A. H. 1905. A Color Notation, GH Ellis Co.
Pavlou, P. A., and Gefen, D. 2004. "Building Effective Online Marketplaces with Institution-Based Trust.," Information Systems Research (15:1), March, pp. 37-59.
Pelet, J. 2010. "Effets de la couleur des sites web marchands sur la mémorisa-tion et sur l'intention d'achat," Systèmes d'Information et Management (15:1), February, .
Pelet, J.-E. and Papadopoulou, P. (2010). The Effect of E-Commerce Websites Colors on Consumer Trust. International Journal of E-Business Research, forthcoming.
Pokorny, J., and Smith, V. 1986. "Colorimetry and Color Discrimination," Handbook of Perception and Human Performance, .
Raykov, T. 1998. "Coefficient Alpha and Composite Reliability With Interrelated Nonhomogeneous Items," Applied Psychological Measurement (22:4), December, pp. 375-385.
Ringle, C., Wende, W., and Will, A. 2005. SmartPLS 2.0 M3 (beta), Hamburg, Germany: University of Hamburg.
Sia, C. L., Lim, K. H., Leung, K., Lee, M. K., Huang, W. W., and Benbasat, I. 2009. "Web Strategies to Promote Internet Shopping: Is Cultural-Customization Needed?," MIS Quarterly (33:3), pp. 491-512.
Stone-Romero, E. F., and Rosopa, P. J., 2004. "Inference Problems with Hierarchical Multiple Regression-Based Tests of Mediating Effects" Research in Personnel and Human Resources Management (23) pp 249-290.
Stone-Romero, E. F., and Rosopa, P. J., 2008. " The Relative Validity of Inferences About Mediation as a Function of Research Design Characteristics" Organizational Research Methods (11:2).
Thayer, R. E. 1986. "Activation-Deactivation Adjective Check List: Current Overview and Structural Analysis," Psychological Reports (58:2), pp. 607-614.
The Standish Group. 1995. "The Standish Group Report: Chaos," , .
Tractinsky, N., Cokhavi, A., Kirschenbaum, M., and Sharfi, T. 2006. "Evaluating the Consistency of Immediate Aesthetic Perceptions of Web Pages," International Journal of Human-Computer Studies (64:11), November, pp. 1071-1083.
Valdez, P., and Mehrabian, A. 1994. "Effects of Color on Emotions," Journal of Experimental Psychology: General (123:4), December, pp. 394-409.
Waggoner, T. 2005. Ishihara Compatible Pseudoisochromatic Plate (PIPIC) Color Vision Test, Good-Lite Company.
Walters, J., Apter, M. J., and Svebak, S. 1982. "Color preference, arousal, and the theory of psychological reversals," Motivation and Emotion (6:3), pp. 193-215.
Wang, Y. D., and Emurian, H. H. 2005. "An Overview of Online Trust: Concepts, Elements, and Implications," Computers in Human Behavior (21:1), pp. 105-125.
Yang, Y., Hu, Y., and Chen, J. 2005. "A Web Trust-Inducing Model for E-Commerce and Empirical Research," in Proceedings of the 7th International Conference on Electronic Commerce, Xi'an, China: ACM, pp. 188-194.