The Influence of Internet Experience on the Judgment of Mobile Web Design

Kiemute Oyibo
University of Saskatchewan
Saskatoon, Canada
kiemute.oyibo@usask.ca

Ifeoma Adaji
University of Saskatchewan
Saskatoon, Canada
ifeoma.adaji@usask.ca

Julita Vassileva
University of Saskatchewan
Saskatoon, Canada
jiv@cs.usask.ca

Web design is a key factor in the success of e-commerce websites. Research has shown that it can be leveraged to achieve positive outcomes in online transaction and user behavior. However, there is limited knowledge on how internet experience influences the perception of mobile web design and user preferences. To bridge this gap, we conducted an empirical study among 115 subjects from China to investigate how internet experience influences the judgement of mobile websites and user preferences. Our results show that users with higher internet experience are more critical of mobile web design than those with lower internet experience. Moreover, users with lower internet experience are more likely to prefer colorful web design than those with higher internet experience. On the other hand, users with higher internet experience are more likely to prefer minimalistic web design than those with lower internet experience. We discuss the implications of our findings in the context of mobile website design.

Web design, mobile website, aesthetics, usability, credibility, visual, color, layout, internet experience, perception, China.

1. INTRODUCTION

Smartphones have become ubiquitous globally thanks to the mobile technology and social media revolutions. For example, between 2005 and 2013 alone, mobile usage grew from 2 billion to 6.8 billion, almost equaling the global population figure of 7 billion (Internet World Stats n.d.). Over the years, the use of the Internet has grown tremendously, with Asia, Europe, Africa and North America accounting for about 48.7%, 17.0%, 10.9% and 8.3%, respectively, of Internet users globally (Internet World Stats 2018). Moreover, with the inception of globalization, mobile commerce websites have become popular such that a website designed in a given part of the world can be accessed by anybody anywhere around the world at any time (Oyibo & Hamada 2013). However, research has shown that the global Internet user population is heterogeneous (Perdue 2001), with different user groups having different preferences and perceptions of mobile websites. As a result, for mobile commerce websites to be successful, there is a need to tailor website design to the different user groups that characterize the Internet. Web design has been recognized as a key factor in the success and survival of e-commerce vendors in the World Wide Web. Web design refers to the process of creating a web artefact (e.g., website) which is purposeful, aesthetic, useful, usable and credible (Flavian et al. 2009). Research has shown that various user characteristics influence the perception and judgment of web designs (Fogg 2003). However, there is limited research in the human-computer interaction (HCI) showing how internet experience (IE) in particular influences user judgment of and preferences for mobile web designs. According to Perdue (2001), "*rather than examining user demographics [in general], it is increasingly important to focus on user experience, including Internet experience*" (p. 27). Specifically, investigating the influence of IE on user judgment of mobile web designs will help designers in tailoring the user interface (UI) of mobile websites to different user groups with different levels of IE. This has the potential of increasing the *visual appeal* and *credibility perception* of mobile websites among the different user groups segmented by IE.

To advance research in this area, we conducted an empirical study to investigate how IE moderates the judgment of mobile websites using Chinese subjects (n = 115) as a case study. We focused on China because it has the largest population of Internet users in the Asian continent, which accounts for about 50% of Internet users worldwide (Internet World Stats 2018). Specifically, over 95% of Internet users in China use the mobile device to access information online, including e-commerce websites (Cheung 2017). This makes it important for HCI researchers to understand how users from this region of the world with different levels of IE perceive and judge mobile websites.

Published by BCS Learning and Development Ltd.
Proceedings of British HCI 2018. Belfast, UK.

To understand user judgment of and preferences for mobile websites based on IE, we based our study on four systematically designed tourist websites, differentiated by UI design features such as layout, color and visual (image). The different combinations of all three criteria resulted in different web designs, out of which we focused on four. The four designs include *multicolor-list-based non-visual website (A)*, *minimalist list-based visual website (B)*, *medium-color list-based visual website (C)* and *medium-color grid-based visual website (D)*. By "visual," we refer to the use of icons or images in the presentation of content in the homepage of a mobile website.

Prior research shows that, when users have no knowledge of the source of a website and the source's characteristics, UI features such as information design and design look are used to determine the *credibility* of a website (Fogg et al. 2002; Hong 2006). Thus, in our study, we asked participants in an online survey to rate and comment on all four designs of mobile websites in terms of *perceived aesthetics*, *usability* and *credibility* based on their first impression. Moreover, we asked them to rank all four web designs from "*most credible*" to "*least credible*" and provide comments on their choice of most credible. To analyze our data, we carried out repeated measure analysis of variance (ANOVA) with respect to each web design attribute to uncover how users with high experience (HE) on the Internet differ from users with low experience (LE). IE is operationalized as the number of years a user has been on or using the Internet.

The results of our study show that there is an interaction effect between IE and web design as well as main effects of both factors. First, with respect to *aesthetics rating*, there is an interaction effect between IE and web design, with the LE group rating web designs A and C higher than the HE group, and both groups not significantly different with respect to web designs B and D. Second, with respect to *usability rating*, there is a main effect of web design, with the general population rating web design D higher than web design A. Third, with respect to *credibility rating* there is a main effect of web design, with the global population rating web design D higher than A, B and C. Finally, with respect to *credibility ranking*, there is a marginal interaction between IE and web design. The HE group ranked B as more credible than A, but there is no significant difference between the rankings of the LE group. On the other hand, the LE group ranked C as more credible than B, but there is no significant difference between the rankings of the HE group. We discuss the implications of our findings and made design recommendations based on our findings.

The rest of the paper is organized as follows. Section 2 and Section 3 focus on the background and related work, respectively. Section 4 and Section 5 focus on research method and results, respectively. Finally, Section 6 and Section 7 dwell on the discussion and conclusion, respectively.

2. BACKGROUND

In this section, we provide an overview of the concepts of *aesthetics*, *usability* and *credibility*. These three web design attributes are widely recognized and studied in the literature (Fogg et al. 2002; Hong 2006; Oyibo & Vassileva 2017).

2.1 Aesthetics

Aesthetics is a concept which is associated with the idea of beauty and visual perception. Historically, it is defined as the branch of philosophy that focuses on "*the nature and appreciation of art, beauty and good taste*" (Mastin 2008). Over the centuries, the concept of *aesthetics* has taken and continues to take on different meanings in different fields, such as art, architecture, design, culture, etc. Specifically, in the eighteenth century, *aesthetics* became known as "*sensory pleasure and delight*" (Hekkert 2006). The concept of *aesthetics* has been theorized as either objective, subjective or an interaction of both views. By "objective," beauty is referred to as a property of the object under evaluation. By "subjective," beauty is conceived as being in the eye of the beholder. In the field of HCI, *visual aesthetics* refers to the pleasing appearance or visual appeal of a HCI artifact (Tractinsky 2002), which could be objective, subjective or both. While we may agree with the interactionist view of *aesthetics*, our study specifically focuses on the subjective view of *aesthetics* as held by Kant (Ginsborg 2014).

Further, research has shown that *aesthetics* is a multidimensional construct. Of the different theories on the multidimensionality of *aesthetics*, Lavie and Tractinsky's (2004) two-dimensional proposition has gained wide acceptance and usage. Based on factorial analysis, the authors found that *visual aesthetics* comprises *classical aesthetics* and *expressive aesthetics*. *Classical aesthetics*, which captures the historical notion of *aesthetics*, entails orderliness, symmetry, proportion, etc. Thus, it is described by terms such as "well-organized," "symmetrical," "clean," "etc. On the other hand, *expressive aesthetics* entails the creativity and originality of designers, including their ability to go beyond conventional standards in the enrichment of their creation. Thus, *expressive aesthetics* is described by terms such as "creative," "fascinating," "sophisticated," "special effect," etc. While *classical aesthetics* basically provides pleasure, *expressive aesthetics* fosters users' involvement with UIs (Lavie & Tractinsky 2004).

2.2 Usability

Usability refers to how easy it is to use a HCI system when users interact with it to achieve their goals. In HCI design, the ultimate goal of *usability* is to: (1) increase users' familiarity and competence in using the system (2) facilitate users' achievement of their objectives while using the system; and (3) enhance users' ability to recall how to use the system in the future, e.g., remembering the location of widgets (Anon n.d.).

2.2 Credibility

Credibility refers to how believable a HCI system is. In other words, it is a measure of how much users believe a system has been developed by credible people (experts) and is trustworthy as a result. *Credibility* is regarded as a perceived quality rather than a property resident in the perceived object (Fogg 2003). Research has shown that credible systems are more likely to be used than less credible systems (Alsudani & Casey 2009). Moreover, *perceived aesthetics* and *usability* have been shown to be good determinants of *perceived credibility* (Oyibo & Vassileva 2016). This means that a HCI system that rates high in *perceived aesthetics* and *usability* is more likely to rate high in *perceived credibility* than the one that rates low in either or both attributes.

3. RELATED WORK

A certain amount of research work has been done on user perception and judgment of websites, especially in the desktop domain. We review a cross-section of them. In the desktop domain, Fogg et al. (2002) carried out a study on what predominantly determines *web credibility*. They found that design look and information design are the strongest determinants of *web credibility*. Hong (2006) conducted a similar study and found that message features, such as quotations and information currency, influence the *credibility perception* of medical websites. However, the author found that structural features, such as domain name, banner and advertisement, do not influence *web credibility*. Sun (2001) investigated the influence of culture on the evaluation of website. The author found that low-context (LC) cultures, e.g., Germany, prefer verbal and hierarchical websites, in which information is organized logically, whereas high-context (HC) cultures, e.g., China, prefer highly visual and colorful websites. Oyibo et al. (2016a) conducted a similar study among Nigerians (HC culture) and Canadians (LC culture) in the mobile domain. They found that Canadians are more critical of mobile websites than Nigerians. Moreover, they found that Nigerians are more attracted to colorful websites than Canadians. They attributed the differences in the judgment of both groups to IE, which is higher for the Canadian group than the Nigerian group. Moreover, Perdue (2001) conducted a study on the influence of behavioral experience on web characteristics (e.g., visual attractiveness and site navigation). Though most of the influences are not significant, they found that, overall, consumer experience negatively influences the assessment of tourism web characteristics such as site navigation. Furthermore, Ahmad et al. (2010) conducted a study on the impact of user experience on the judgment of *web credibility*. They found that novice users made their *credibility judgment* based on mostly visual appearance and structural features, whereas expert users made their *judgment of credibility* based on a more thoughtful assessment of the various aspects of structural and message features of a website.

In the foregoing related work review, we found that there are some gaps with respect to the influence of IE on user judgment of websites. Perdue's (2001) study did not investigate IE, in particular, but consumer experience, in general, which comprised IE and behavioral experience (e.g., skiing and tourist experience). Moreover, their study focused on Western population. Similarly, Ahmad et al.'s (2010) study only focused on 15 native English and non-heritage Spanish speakers, which may not generalize to other populations. However, our study investigates how IE influences the judgment of websites using Chinese subjects as a case study. Moreover, it focuses on the mobile domain, which has been scarcely investigated.

4. METHOD

In this section, we present our research objective and design, research hypotheses, measurement instruments and participants' demographics.

4.1 Research Objective and Design

Despite the theory that user characteristics influence the judgment of websites (Fogg 2003), little has been done with respect to the influence of IE in the mobile domain (Ahmad et al. 2010). In this paper, using a mixed-method approach, we aim to answer the following research questions:

(i) *How does Internet experience influence user judgment of mobile websites?*
(ii) *Can the influence of Internet experience on the user judgment of mobile websites be replicated across different website designs?*

To address the above research questions, we came up with four different website designs (Figure 1) based on three design characteristics: *layout*, *color* and *visual*, which Sun (2001) regarded as cultural markers. The choice of these four designs is explained in Oyibo & Vassileva (2017b). Meanwhile, with respect to layout, we considered two levels: grid and list. Secondly, with respect to color scheme, we considered three levels: *minimalistic*, *medium-color* and *multicolor*.

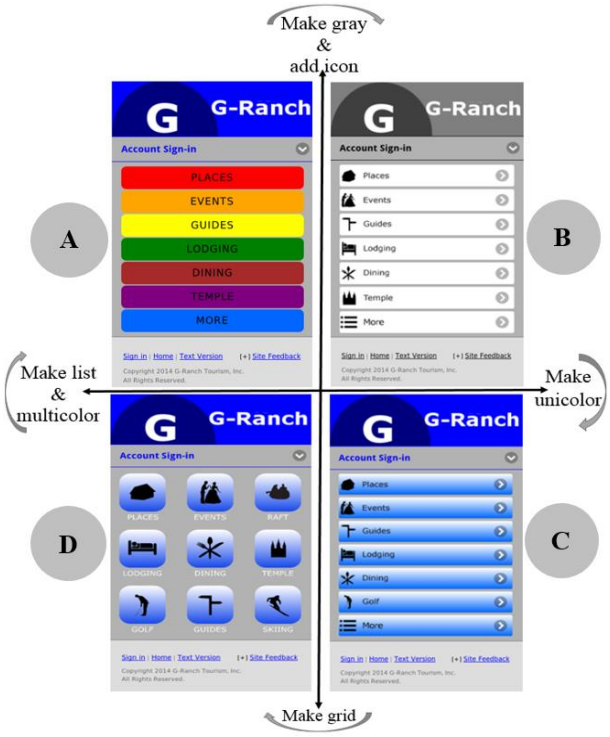

Figure 1: *Mobile website designs evaluated by each subject in the order—C, A, B and D* (Oyibo et al. 2016a)

Finally, with respect to visual, we considered two levels: "*with image/icon*" and "*without image/icon.*" However, in our user study, to prevent participant fatigue, we were not able to investigate all of the UI configurations resulting from the different combinations of the different levels of layout, color and visual. As a result, we only investigated four web designs, which are described as follows:

(i) Multicolor list-style non-visual design (A)
(ii) Minimalist list-style visual design (B)
(iii) Medium-color list-style visual design (C)
(iv) Medium-color grid-style visual design (D)

4.2 Research Hypotheses

Based on existing findings in the literature, we formulated three hypotheses with respect to the influence of IE on user judgment of mobile websites. Specifically, in the study of tourist websites, Perdue (2001) found a number of negative correlations between consumer experience and website evaluations. Therefore, in our study, we hypothesize as follows:

H1: HE users will be more critical of the *perceived aesthetics* of the four website designs than LE users.

H2: HE users will be more critical of the *perceived usability* of the four website designs than LE users.

H3: HE users will be more critical of the *perceived credibility* of the four website designs than LE users.

Moreover, given the paucity of research in this topic area, we used an exploratory approach to investigate which of the four designs will be judged the best and worst by the different IE groups.

4.3 Measurement Instruments

In measuring the three web design attributes, we used existing validated instruments from the literature. To measure *aesthetics*, we used the adapted short-version scale (van Schaik & Ling 2009) proposed by Lavie and Tractinsky (2004). The scale is composed of two dimensions: *classical aesthetics* and *expressive aesthetics*, which we assumed to be unidimensional for the purpose of this study. Second, to measure *usability*, we used Lavie and Tractinsky's (2004) 5-item scale. Finally, we used a sing-item scale—which has been found reliable as a multi-item scale (Bergkvist & Rossiter 2007)—to measure *credibility*. The *aesthetics* and *usability* scales ranged from "*Strongly Disagree (1)*" to "*Strongly Agree (7)*," while the *credibility* scale ranged from "*Very Bad (1)*" to "*Very Good (7)*," Table 1 shows all three scales together with their respective items. All of the items in the *aesthetics* and *usability* scales, in respect of each website design, were randomized in a block and presented to participants. Next, the questions on *credibility* and commenting on the design were presented. After answering questions on all four designs in the order—C, A, B and D—the question on the *credibility ranking* of the four web designs was presented to participants. In addition, we asked participants to provide the reason they ranked one of the four web designs as the most credible.

Table 1: *Measurement scales for web design attributes*

Measure	Items in Scale
Aesthetics Rating	*Classical Aesthetics* 1. Visual; 2. Clean; 3. Pleasant *Expressive Aesthetics* 1. Creative; 2. Fascinating; 3. Sophisticated
Usability Rating	1. Clear design. 2. Convenient to use. 3. Easy to navigate. 4. Easy orientation. 5. Easy to use
Credibility Rating	Credibility level
Credibility Ranking	Please give the reason behind your choice of most credible.

4.4 Participants

Our study was approved by our university's behavioral research ethics board. Participants in a Chinese university and a Canadian university were invited via email and the latter university's website, respectively, to participate anonymously in the study. Table 2 shows the demographics of the subjects (n = 115). 105 were resident in China, while 10 were resident in Canada. 71.3% had less than 10 years of IE (LE users), while 28.7% had 10 years and above of IE (HE users).

Table 2: *Participants' demographics*

Criteria	Group	HE	LE	All	Perc.
Gender	Male	14	32	46	40.0
	Female	18	45	63	54.8
	Unidentified	1	5	6	5.2
Age	18-24	11	40	51	44.3
	>24	22	42	64	55.7
Educ. Qual.	High School	16	60	76	66.1
	Bachelor	12	13	25	21.7
	Others	5	9	14	12.2
Internet Access	Always	19	29	48	41.7
	Often	14	52	66	57.4
	Rarely	0	1	1	0.9
Subtotal (No.)		33	82	115	100
Percent (%)		28.7	71.3	115	100

LE = Low Internet Experience (<10 years)
HE = High Internet Experience (>= 10 years)

4.5 Data Analysis

We began our analysis by testing the normality of our data. The result of the test showed that the data is not normal. Thus, we resorted to conducting a non-parametric reliability test for the web design constructs and non-parametric ANOVA.

5. RESULTS

In this section, we present the result of the reliability test, the mean performance (scores) of all four measures and the ANOVA.

5.1 Reliability Analysis

We conducted McDonald's omega (ω) reliability test (Dunn et al. 2014) on the *aesthetics* and *usability* constructs using the "psych" package in R. The results for all four web designs met the recommended reliability requirement (ω >= 0.7).

5.2 Mean Rating and Ranking of Measures

To determine how each of the two IE groups perceived the four website designs, we computed the overall average score of each of the four measures (*aesthetics rating*, *usability rating*, *credibility rating* and *credibility ranking*) as shown in Figure 2. Overall, especially with respect to websites A and C, the LE group rated the *aesthetics* and *credibility* attributes higher than the HE group. For example, with respect to website A, the average *aesthetics* and *credibility ratings* of the LE group are 4.34 and 4.35, respectively, while those of the HE group are 3.60 and 3.80, respectively. Moreover, with respect to the ranking of the website designs, the LE group ranked website A (2.21) way higher than the HE group (1.76). However, the HE group ranked websites B (2.52) way higher than the LE group (2.22). Repeated measure ANOVA will determine whether there is an interaction between IE and website design and/or a main effect of both factors. We carried out these analyses in the subsequent subsections.

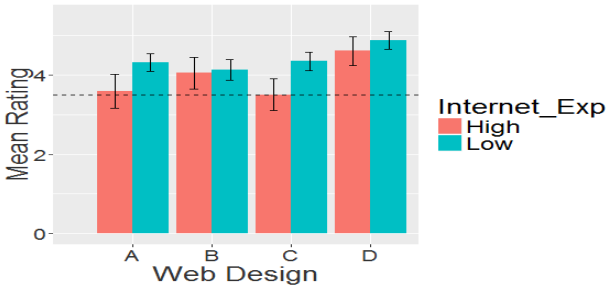

Figure 2a: *Aesthetics rating of all web designs*

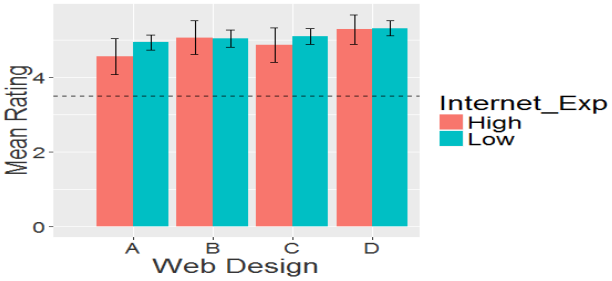

Figure 2b. *Usability rating of all web designs*

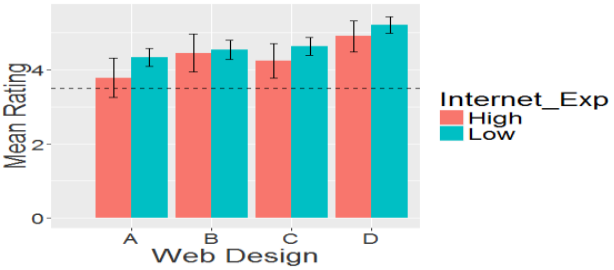

Figure 2c. *Credibility rating of all web designs*

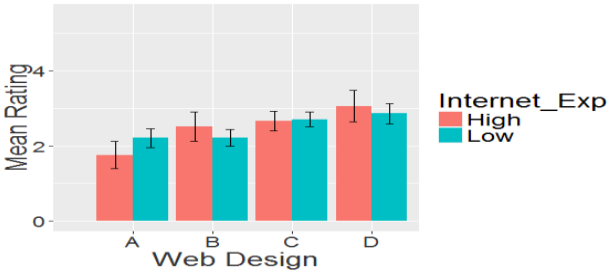

Figure 2d. *Credibility ranking of all web designs*

5.3 Two-Way Mixed ANOVA

We carried out a two-way mixed ANOVA (Wobbrock et al. 2011) using the "ARTool" package (Kay & Wobbrock 2016) in R to determine the main and interaction effects of IE and web design. Table 3 shows the results. With respect to *aesthetics rating*, there is an interaction between IE and web design ($F_{3,\ 452}$ = 2.75, $p < 0.05$). Secondly, with respect to *usability rating*, there is a main effect of web design ($F_{3,\ 452}$ = 3.46, $p < 0.05$). Thirdly, with respect to *credibility rating*, there is a main effect of IE ($F_{1,\ 452}$ = 4.28, $p < 0.05$) and web design ($F_{3,\ 452}$ = 13.35, $p < 0.001$). Finally, with respect to *credibility ranking*, there is a main effect of web design ($F_{3,\ 452}$ = 9.78, $p < 0.001$) and a marginal interaction effect ($F_{3,\ 452}$ = 2.21, $p < 0.086$).

Table 3: *Main and interaction effects between internet experience and website design*

Measure	Interaction Effect	Internet Experience	Web design
Aesthetics Rating	($F_{3, 452}$ = 2.75, **p < 0.05)**	($F_{1, 452}$ = 17.35, **p < 0.001)**	($F_{3, 452}$ = 9.53, **p < 0.001)**
Usability Rating	$F_{3, 452}$ = 0.95, p = n.s);	($F_{1, 452}$ = 0.09, p = n.s)	($F_{3, 452}$ = 3.46, **p < 0.05)**
Credibility Rating	($F_{3, 452}$ = 0.37, p = n.s)	($F_{1, 452}$ = 4.28, **p < 0.05)**	($F_{3, 452}$ = 13.35, **p < 0.001)**
Credibility Ranking	($F_{3, 452}$ = 2.21, p = 0.086)	($F_{1, 452}$ = 0.23, p = n.s	($F_{3, 452}$ = 9.78, **p < 0.001**

5.4 IE Effect: Between Group Analyses

We carried out a between-group effect analysis [Kruskal-Wallis rank sum test (Pohlert 2014)] on the four measures to further determine in finer details how the HE and LE groups differ in the judgment (rating and ranking) of the web designs. The result (see Table 4) shows that there is a simple main effect of IE ($p < 0.01$) on the *aesthetics ratings* of web designs A and C. Specifically, the LE group rated both designs higher than the HE group. However, both groups do not significantly differ ($p > 0.05$) with respect to web designs B and D. Furthermore, with respect to *usability rating* and *credibility* rating, there is no simple main effect of IE for any of the web designs. However, with respect to *credibility ranking*, there is a simple main effect of IE ($p < 0.05$) with respect to web design A, with the LE group ranking it higher than the HE group.

5.5 Web Design Effect: Within Group Analyses

We also carried out a within-group post-hoc pairwise comparisons to determine the simple main effect of web design at the global and subgroup levels. Tukey method was used to correct for familywise error. Our results (see Table 5) show that there are simple main effects of web design for a number of the measures at the global and subgroup levels. In particular, with respect to *aesthetics rating, credibility rating* and *credibility ranking*, there is a simple main effect of web design, as evident in the A-D, B-D and C-D comparisons, most of which are significant ($p < 0.05$) at the global and subgroup levels, e.g., A-D. Moreover, there is an interaction between IE and web design with respect to some of the pairwise comparisons: B-D in *aesthetics rating*, B-D and C-D in *credibility rating*, and A-B, B-C and B-D in *credibility ranking*. However, the two-way omnibus ANOVA in subsection 5.3 shows that it is only the interaction with respect to *aesthetics rating* (B-D) that is significant ($p < 0.05$). The interactions with respect to *credibility rating* are non-significant ($p > 0.05$), while those with respect to *credibility ranking* are marginally significant ($p = 0.086$).

Table 4: *Between-group comparisons of web designs*

Measure	HE Group	LE Group	P-Value
Aesthetics Rating			
A	**3.60**	**4.33**	**P < 0.01**
B	4.06	4.13	p = n.s
C	**3.51**	**4.35**	**p < 0.001**
D	4.61	4.88	p = n.s
Overall Av.	**4.94**	**4.42**	**p < 0.001**
Usability Rating			
A	4.55	4.93	p = n.s
B	5.05	5.03	p = n.s
C	4.87	5.09	p = n.s
D	5.28	5.31	p = n.s
Overall Av.	4.94	5.09	p = n.s
Credibility Rating			
A	3.79	4.34	p = n.s
B	4.45	4.54	p = n.s
C	4.24	4.63	p = n.s
D	4.91	5.21	p = n.s
Overall Av.	**4.35**	**4.68**	**p < 0.05**
Credibility Ranking			
A	**1.76**	**2.21**	**p < 0.05**
B	2.52	2.22	p = n.s
C	2.67	2.71	p = n.s
D	3.06	2.86	p = n.s
Overall Av.	2.50	2.50	p = n.s

LE = Low Internet Experience (<10 years)
HE = High Internet Experience (>= 10 years)

Table 5: *Within-group comparisons of web designs*

Comp.	Global	HE Group	LE Group
Aesthetics Rating			
A-B	p = n.s	p = n.s	p = n.s
A-C	p = n.s	p = n.s	p = n.s
A-D	**p < 0.001**	**p < 0.001**	**p < 0.001**
B-C	p = n.s	p = n.s	p = n.s
B-D	**p < 0.001**	p = n.s	**p < 0.001**
C-D	**p < 0.001**	**p < 0.001**	**p < 0.001**
Usability Rating			
A-B	p = n.s	p = n.s	p = n.s
A-C	p = n.s	p = n.s	p = n.s
A-D	**p < 0.001**	p < 0.067	p < 0.051
B-C	p = n.s	p = n.s	p = n.s
B-D	p = n.s	p = n.s	p = n.s
C-D	p = n.s	p = n.s	p = n.s
Credibility Rating			
A-B	**p < 0.01**	p = n.s	p = n.s
A-C	p = 0.067	p = n.s	p = n.s
A-D	**p < 0.001**	**p < 0.001**	**p < 0.001**
B-C	p = n.s	p = n.s	p = n.s
B-D	**p < 0.05**	p = n.s	**p < 0.001**
C-D	**p < 0.01**	p = n.s	**p < 0.01**
Credibility Ranking			
A-B	p = n.s	**p < 0.05**	p = n.s
A-C	**p < 0.001**	**p = 0.001**	**p < 0.05**
A-D	**p < 0.001**	**p < 0.001**	**p < 0.001**
B-C	p = 0.08	p = n.s	p = 0.05
B-D	**p < 0.01**	p = n.s	**p < 0.001**
C-D	p = n.s	p = n.s	p = n.s

LE = Low Internet Experience (<10 years)
HE = High Internet Experience (>= 10 years)

5.6 Sentiment Analysis

Figure 3 shows the sentiment analysis[1] of participants' comments on all four web designs. It captures positive (POS), neutral (NEU) and negative (NEG) sentiments. Overall, web designs A and B elicited the highest percent of negative comments, while web design D—followed by web design C—elicited the highest percent of positive comments. For example, regarding A and B, 42.5% of the global comments on either of them is negative, while, regarding D and C, only 11.32% and 30.2% of the global comments on them, respectively, are negative. Overall, the HE group is more critical of the web designs than the LE group, especially with respect to A and C. For example, regarding A, 57.7% of the HE group's comments are negative, while only 37.5% of the LE group's comments are negative. Similarly, regarding web design C, 50.0% of the HE group's comments are negative, while only 23.8% of the LE group's comments are negative.

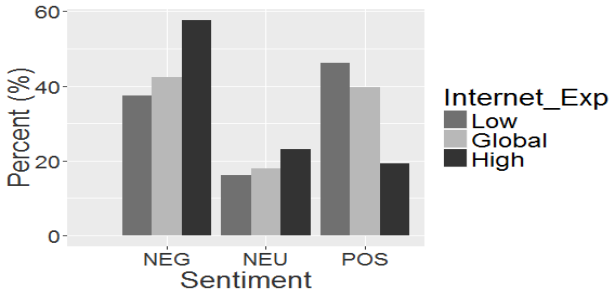

Figure 3a. *Sentiment polarity for web design A (NEG = Negative, NEU = Neutral, POS = Positive)*

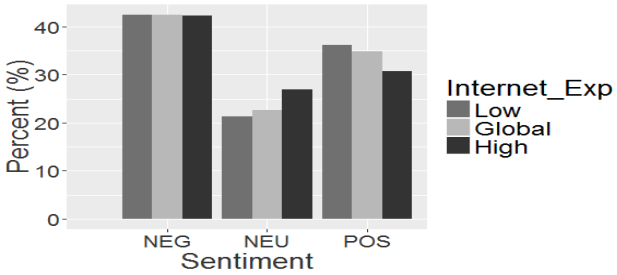

Figure 3b. *Sentiment polarity for web design B (NEG = Negative, NEU = Neutral, POS = Positive)*

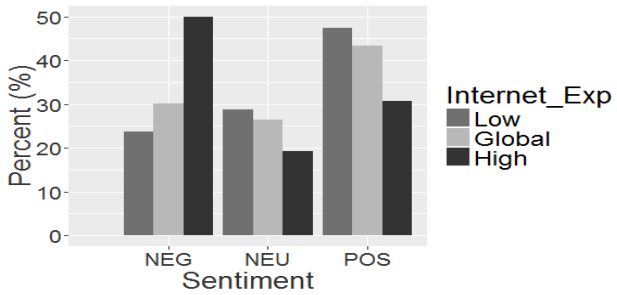

Figure 3c. *Sentiment polarity for web design C (NEG = Negative, NEU = Neutral, POS = Positive)*

[1] The text for the sentiment analysis was translated by a Chinese Researcher from Chinese to English. We acknowledge that the translation may not have been perfect.

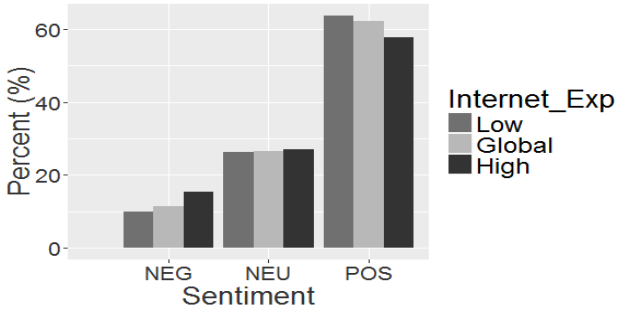

Figure 3d. *Sentiment polarity for web design D (NEG = Negative, NEU = Neutral, POS = Positive)*

6. DISCUSSION

We have presented the judgment of four website designs by two different groups of users having different levels of IE: high and low. Our results show that both groups are similar as well as different in the judgment of the mobile web designs. Specifically, with respect to the judgment of the minimalist list-based web design (B) and the medium-color grid-based web design (D), there is no significant difference between both groups quantitatively (based on all four measures) and qualitatively (based on the sentiment analysis). However, with respect to the judgment of the multicolor list-based design (A) and the medium-color list-based design (C), both groups significantly differ quantitatively and qualitatively. Overall, the HE group is more critical of the *aesthetics* of both websites than the LE group ($p < 0.01$). Moreover, the LE group ranked the *credibility* of A more favorably than the HE group did ($p < 0.05$). In the rest of this section, we discuss our findings in the context of our hypotheses and their implications in the context of mobile web design.

6.1 Validation of Pre-Formulated Hypothesis

In the light of our research questions and hypotheses, our main findings can be summarized as shown in Table 6. In addition to the supported and unsupported hypotheses, we have included our secondary findings (H4 and H5), which resulted from our exploratory analysis. We discuss all of the findings in the subsequent subsections.

6.1.1. Validation of First Hypothesis

Our first hypothesis (H1) states that HE users will be more critical of the *perceived aesthetics* of the four web designs than LE users. The result of our analysis shows that our hypothesis is partly supported. Overall, HE users were more critical of the web designs. Specifically, the HE users were more critical of the multicolor list-style non-visual design (A) and the medium-color list-style visual design (C) than LE users (see Table 4). However, there is no significant difference between both groups with respect to the minimalist list-style visual design (B) and the medium-color grid-style visual design (D). Interestingly, our qualitative analysis of participants' comments on web designs

Table 6: *Summary of main findings*

H	Hypotheses	Remark
1	HE users will be more critical of the *perceived aesthetics* of the four website designs than LE users.	Supported overall and for A and C, but not for B and D.
2	HE users will be more critical of the *perceived usability* of the four website designs than LE users.	Not supported overall and at each level of web design
3	HE users will be more critical of the *perceived credibility* of the four website designs than LE users.	Supported overall, but not at each level of web design
4*	LE users will rank the *credibility* of the multicolor website design (A) higher than HE users.	Supported
5*	HE users will rank the *credibility* of the minimalist website design (B) higher than LE users.	True but not significant

*Hypotheses were not pre-stated prior to analysis; these specific findings are based on exploratory analysis.

A and C confirms the quantitative finding (i.e., the significant difference between the LE group and the HE group in the *aesthetics ratings* of both designs). The sentiment analysis with respect to both designs (Figure 3a and Figure 3c) shows that the HE group is more critical of the website designs than the LE group. With respect to web design A, for example, the percent of positive comments of the LE group (46.3%) is higher than that of the HE group (19.2%), while the percent of the negative comments of the HE group (57.7%) is higher than that of the LE group (37.5%). Similarly, with respect to design C, the percent of positive comments of the LE group (47.5%) is higher than that of the HE group (30.8%), while the percent of negative comments of the HE group (50.0%) is higher than that of the LE group (23.8%).

In sum, qualitatively and qualitatively, HE users were more critical of the *perceived aesthetics* of web designs A and C than LE users. Therefore, our first hypothesis (H1) is supported and replicated across two different levels of the web designs (A and C). The first hypothesis is also confirmed by the qualitative results. One possible reason why the result of the sentiment analysis confirms that of the between-group analysis based on *aesthetics*, and not that based on *usability*, for example, is that a large part of participants' judgment of a website is focused on design look (e.g., color scheme, information layout, etc.), which is related to *visual aesthetics* (see example comments in subsection 6.2). Thus, in our study, we see that the differences between both IE groups with respect to the *aesthetics ratings* of the four designs are reflected in the differences in the polarity of the sentiments they expressed towards the four web designs.

6.1.2. Validation of Second Hypothesis

Our second hypothesis (H2) states that HE users will be more critical of the *perceived usability* of the four web designs than LE users. However, the results of our between-group analysis show that there is no significant difference between both groups overall and at each level of the web design. This indicates that, irrespective of the level of IE, both groups perceived and judged all of the four web designs in a similar way. In other words, there is no significant difference between both groups' *usability ratings* of each web design. This also suggests that the level of IE of users does not determine the level of the *perceived usability* of a UI. One plausible explanation for this finding of no significant difference between both groups is that participant care less about *perceived usability* in the judgment of website (e.g., *credibility*) than they do about hedonic attributes such as *aesthetics* (Kurosu et al. 1995; Oyibo & Vassileva 2017; Oyibo & Vassileva 2016). Another plausible explanation is that participants did not actually use the different web designs. As a result, their judgment was based on *perceived usability* rather than *actual usability*, which judgment may be influenced by IE. However, this finding (no effect of IE on *perceived usability*) needs to be further investigated in future work.

6.1.3. Validation of Third Hypothesis

Our third hypothesis (H3) states that HE users will be more critical of the *perceived credibility* of the four web designs than LE users. Our between-group analysis results show that there is no significant difference between both groups at each level of the web design despite the fact that the ratings of all four designs are consistently higher for the LE group than the HE group, especially for A, C and D. However, while there is no significant difference between both groups at each level of web design, there is a main effect of web design at the global level. In other words, the LE group's overall *credibility rating* of the four designs is significantly higher than the HE group's. This implies that HE users, overall, are more critical of the *credibility* of web designs than LE users. However, there is a need for further investigation (e.g., in other domains, among other demographics, etc.) to confirm our finding.

6.2 Findings Based on Exploratory Analysis

The starred findings (H4 and H5) in Table 6 are based on the exploratory analysis of the *credibility ranking* of the four web designs by the two different IE groups. We discuss them, together with the results of our sentiment analysis, in this subsection.

6.2.1. Validation of Fourth and Fifth Hypotheses

With respect to the *credibility ranking* of the four web designs by both IE groups, we found that there is a significant difference between them regarding the multicolor web design (A). The LE group significantly ranked A higher than the HE group

(see Table 5). Therefore, the fourth hypothesis (H4) supported. This suggests that the LE group is more favorable towards colorful websites than the HE group. In contrast, the HE group seems to be more favorable towards the minimalist design (B) than the LE group. The HE group (2.52) ranked B higher than the LE group (2.22). However, the difference in ranking is not statistically significant ($p > 0.05$). Thus, the fifth hypothesis (H5) is not supported.

6.2.2. Most Preferred Web Design by Both Groups
Overall, irrespective of the level of IE, web design D is significantly preferred to the other three web designs (A, B and C). For example, as shown in Table 5, based on *aesthetics rating* at the global and subgroup levels, web design D—the highest-rated design—is significantly different ($p < 0.001$) from the other designs, except for B in the HE group. This finding is further confirmed in the within-group pairwise comparisons for the global population and LE group (see Table 5), in which the *credibility rating* of D is significantly higher ($p < 0.01$) than that of A, B and C. Similarly, the pairwise comparisons based on *credibility ranking* for the global population and LE group shows that the ranking of D is significantly higher ($p < 0.01$) than that of A and B. Thus, overall, web design D is judged the best by both groups. This is confirmed by the qualitative results, in which D received more positive comments (63.8% and 57.7%) and less negative comments (10.0% and 15.4%) from both LE and HE groups, respectively, than any of the other three web designs (A, B and C). The following shows some of the positive comments from participants regarding web design D:

(i) "*Design D, similar to [grid] styles in phones. Looks more trustful*" [P7, HE].
(ii) "*Closer to usage habit (like apps arrangement in phones*" [P2, HE].
(iii) "*Beautiful interface, easy to use*" [P70, LE].
(iv) "*Layout good, color comfortable, icons vivid*" [P3, HE].
(v) "*Easy to understand, accompanied with icons*" [P36, LE].

As seen in the above comments, the most common reasons why participants judged web design D as the most aesthetic and credible include "easy to understand and use," "has large icons," and "has a layout similar to the smartphone's," which they were already familiar with. It is interesting to note that, although, based on the quantitative result for *usability rating* (see Table 5), there is no significant difference between D and the other web designs (except A), the sentiment analysis reveals that *perceived ease of use* (stemming from the layout of the content) is one of the main reasons why D is the most preferred by both IE groups. Generally, users view the grid design (D), unlike the list design (C), as less prone to errors when clicking on a given item. This is due to the relatively large spacing of its items (content-icons) (Oyibo et al. 2016b). Compared to C, this unique feature of the grid layout, owing to the halo effect, must have enhanced the perception of other design attributes of D, such as *perceived aesthetics* and *credibility*.

6.2.3. Least Preferred Web Design by IE Groups
Overall, the results of our pairwise comparisons with respect to *aesthetics*, *usability* and *credibility ratings*, show that there is no significant different between each pair of web designs A, B and C at the subgroup level (see Table 5). However, the within-group analysis based on *credibility ranking* reveals the relative dislike of the multicolor design (A) by the HE group and the minimalist design (B) by the LE group. In particular, the HE group significantly ranked the other three designs (including B, which some participants deemed "boring") higher than A. In addition, though not significant, we see that the HE group rated the *usability* and *credibility* of B, C and D higher than A. However, this is not the case with the LE group. Specifically, the LE group did not significantly rank B higher than A like the HE group did (see Table 5). Even, in some cases, the LE group rated A (4.33) as more aesthetic than B (4.13), unlike the HE group that rated B (4.06) as more aesthetic than A (3.60), though the respective differences are not statistically significant ($p > 0.05$). Moreover, the LE group significantly ranked C and D higher than B, which is not the case with the HE group. All of these results are an indication of the following:

(i) The HE group tends to be more favorable towards minimalist designs than the LE group and than highly colorful designs.
(ii) The LE group tends to be more favorable towards highly colorful designs than the HE group.

The first part of the first statement is proven quantitatively by the B-C and B-D *credibility ranking* comparisons mentioned above. In the comparisons (Table 5), C and D are significantly ranked higher than B by the LE group but there is no significant difference between each pair in the HE group. The second part is proven by the A-B *credibility ranking* comparison, in which B is significantly ranked higher than A by the HE group. This is confirmed by the qualitative result. The HE group gave more negative comments on A (57.7%) than on B (42.3%). Moreover, the HE group gave less positive comments on A (19.23%) than on B (30.8%). On the other hand, the second statement is confirmed by the qualitative results. Apart from the LE group rating the *aesthetics* of A higher than the HE group (Table 4), the LE group gave more positive comments (46.3%) and less negative comments (37.5%) on A than the HE group did (19.2% and 57.7%, respectively).

6.2.4. General Remarks on Web Design A
Overall, participants perceived web design A as the least credible (see Figure 2 and Table 5). We now present some of the main reasons why they perceived web design A as the least credible. First and foremost, participants were critical of the color scheme of A due to its multiplicity, complexity, poor coordination and harmony, as evident in the following sample comments.

(i) "*The color is messy. I am not sure if there is special reason for such color arrangement*" [P9, HE].
(ii) "*There are too many colors on the buttons*" [P88. LE].
(iii) "*I am uncomfortable with the color, but can't tell the reason why*" [P25, HE].
(iv) "*The colors are too complex*" [P92, LE].

Secondly, some participants raised concerns regarding the list layout of A; they did not like it. We surmise that their dissatisfaction with the multicolor scheme of A, due to the halo effect, must have affected their judgment of its layout as well. Their displeasure with A's layout are captured in the following comments:

(i) "*Complex color, unreasonable navigation*" [P3, HE].
(ii) "*I don't like the layout design*" [P73. LE].
(iii) "*Too many colors, the layout is not tidy*" [P32, LE].

However, despite the negative comments on web design A, there are positive/neutral comments as well, especially from the LE group. This may not be surprising given that, overall, both groups of users rated A favorably, i.e., above the neutral value of 3.5. Generally, prior research (Sun 2001) has shown that Chinese users like colorful designs. In our study, some participants, especially the LE group, actually preferred the multicolor scheme (A) to the medium-color scheme (C). For example, P69, an LE participant, commented thus: "*Just fine, better than prior one [C]"*. Similarly, P29, an LE participant, commented thus: "*Color is more vivid than prior one [C]"*. These participants, in particular, actually preferred the multicolor design (A) to the medium-color (blue) design (C). The following are some more favorable/neutral comments from participants regarding web design A. Overall, more positive comments were given by the LE group than the HE group as shown in Figure 3a.

(i) "*So so, neither good nor bad, no feeling*" [P13, HE].
(ii) "*Design A is simple and clean, vivid color, more attractive*" [P92, LE].
(iii) "*I like such design, simple and easy to understand*" [P52, LE].
(iv) "*Simple to use*" [107 LE].
(v) "*Like this design*" [P79, LE].

6.2.5. Relatively Favorable View of Web Design B
Given the minimalistic color scheme for web design B, one would have expected that it would be preferred less or judged less favorably than the medium-color (blue) design (C), both of which share the same layout. Overall, the results of our pairwise comparisons (see Table 5) for both groups show that there is no significant difference between B and C with respect to all four measures. The only significant difference ($p < 0.05$) is in the *credibility ranking* of both web designs by the LE group. In particular, for the HE group, there is no significant difference between B and C and even between B and D (the most preferred design) with respect to all four measures. This indicates that the B's minimalist design is perceived as good as C's and D's medium-color design by the HE group in particular. One plausible reason for this favorable view of B by the HE group, as we gathered from our sentiment analysis, is that the minimalist design is seen as simple, neat, professional and official, as seen in the following comments:

(i) "*Like it, simple and clear*" [P73, HE].
(ii) "*Grey looks neat, business style*" [P16, HE].
(iii) "*Simple and neat, with a sense of professionalism*" [P73, LE].

This finding suggests that vendors, e.g., companies, could still adopt the minimalist design without being hurt by users' perception and judgment, especially by highly experienced users, who perceive the minimalistic design as professional and credible. For example, the HE group rated the minimalist design (B) as more aesthetic, usable and credible than the medium-color design (C), though the differences are not significant (see Figure 2d and Table 4).

6.3 Summary of Main Findings, Implications and Contributions
For easy and quick access to the results of this study, we summarize the main findings as follows:

(i) HE users are more critical of mobile web designs than LE users.
(ii) HE users are more likely to prefer minimalistic web designs than highly colorful web designs and than LE users.
(iii) LE users are more likely to prefer highly colorful web designs than HE users.
(iv) Both HE and LE users prefer the grid layout to the list layout.

The implication of these findings is that HCI designers will have to do more *aesthetically* and *professionally* when designing mobile websites for the HE group in China than when designing for the LE group. The HE group tends to be more critical of the *aesthetics* and *credibility* of mobile websites.

Moreover, they tend to prefer more professional website (with less flashy colors) than the LE group. Our main contribution to the body of knowledge is that we have provided empirical evidence on how both HE and LE users differ in the judgment of mobile websites. Specifically, we have shown that HE users are more critical of mobile websites than LE users. The results of our finding can be leveraged by HCI designers in tailoring mobile websites to the individual groups in China to enhance their *perceived aesthetics* and *credibility*.

6.4 Limitations and Future Work

There are a number of limitations in our study. The first limitation is that it is based on user perception and not the actual usage of the investigated tourism website designs in a controlled experimental setting. Thus, our findings may not generalize to an actual context of use in which users have to use the different designs to search for tourist information online. The second limitation of our study is that our findings are based on participants from China (based on residency and/or nationality). As a result, our findings may not generalize to other populations or cultures. For this reason, we recommend that future work should focus on investigating other demographics than China to confirm the generalizability of our findings.

7. CONCLUSION

In this paper, we presented the influence of IE on the judgment of mobile web designs in the tourism domain using Chinese subjects (n = 115) as a case study. Our results reveal that: (i) HE users are more critical of mobile web designs than LE users; (ii) HE users are more likely to prefer minimalistic web designs than LE users; and (iii) LE users are more likely to prefer (highly) colorful web designs than HE users. Overall, the general population prefer a medium-color to a minimalistic or highly colorful web design. In addition, they prefer a grid-based to a list-based web design. Thus, in a one-size-fits-all mobile web design, designers should adopt the middle ground in the use of colors in their user interface design. Their web designs should neither be too minimalistic nor over colorful. Moderate use of colors should be the key to a successful mobile web design. Finally, in terms of *usability*, designers should give priority to the grid-based design over the list-based design in the organization of their content in the homepage of their websites. These findings and recommendations hold a potential for better understanding and meeting the different visual needs and expectations of the different user groups at the level of perception. In future work, we aim to extend our study to other demographics. Specifically, we look forward to investigating the generalization of our findings to mobile website users in some other countries and continents, e.g., Africa, North America, etc.

8. REFERENCES

Ahmad, R. et al., 2010. The impact of user experience levels on web credibility judgments. In *Proceedings of the ASIST Annual Meeting*.

Alsudani, F. & Casey, M., 2009. The Effect of Aesthetics on Web Credibility. *23rd British HCI Group Annual Conference on People and Computers: Celebrating People and Technology*, pp.512–519. Available at: http://epubs.surrey.ac.uk/3035/4/licence.txt.

Anon, Usability: User Experience (UX) Topic Overview/Definition. *International Design Foundation*.

Bergkvist, L. & Rossiter, J.R., 2007. The Predictive Validity of Multiple-Item Versus Single-Item Measures of the Same Constructs. *Journal of Marketing Research*, 44(2), pp.175–184.

Cheung, M.-C., 2017. More than 95% of Internet Users in China Use Mobile Devices to Go Online. *EMarketer*. Available at: https://www.emarketer.com/Article/More-than-95-of-Internet-Users-China-Use-Mobile-Devices-Go-Online/1015155 [Accessed March 15, 2018].

Dunn, T.J., Baguley, T. & Brunsden, V., 2014. From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), pp.399–412.

Flavian, C., Gurrea, R. & Orús, C., 2009. Web design: a key factor for the website success. *Journal of Systems and Information Technology*, 11(2), pp.168–184. Available at: http://www.emeraldinsight.com/doi/10.1108/13287260910955129.

Fogg, B.J. et al., 2002. *How Do People Evaluate a Web Site ' s Credibility ? Results from a Large Study*, Available at: https://dejanseo.com.au/media/pdf/credibility-online.pdf.

Fogg, B.J., 2003. Prominence-interpretation theory: Explaining how people assess credibility online. In *CHI'03 extended abstracts on human factors in computing systems*. pp. 722–723.

Ginsborg, H., 2014. Kant's Aesthetics and Teleology. *The Stanford Encyclopedia of Philosophy*. Available at: http://plato.stanford.edu/archives/fall2014/entries/kant-aesthetics/%5Cninternal-pdf://0.0.0.169/kant-aesthetics.html.

Hekkert, P., 2006. Design aesthetics: principles of pleasure in design. *Psychology science*, 48(2), p.157.

Hong, T., 2006. The influence of structural and message features on web site credibility. *Journal of the American Society for Information Science and Technology*, 57(1), pp.114–127.

Internet World Stats, 2018. Internet Users in the World by Regions - Decemeber 31, 2017. *Miniwatts Marketing Group*. Available at: https://www.internetworldstats.com/stats.htm [Accessed March 15, 2018].

Internet World Stats, Mobile Internet - Mobile Phones and Smart Mobile Phones. *Miniwatts Marketing Group*. Available at: https://www.internetworldstats.com/mobile.htm [Accessed March 15, 2018].

Kay, M. & Wobbrock, J.O., 2016. Package "ARTool." *CRAN Repository*, pp.1–13. Available at: https://cran.r-project.org/web/packages/ARTool/ARTool.pdf [Accessed June 20, 2018].

Kurosu, M., Kashimura, K. & Creativity, O.F., 1995. Apparent usability vs. inherent usability: experimental analysis on the determinants of the apparent usability. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, pp.292–293. Available at: http://portal.acm.org/citation.cfm?id=223680.

Lavie, T. & Tractinsky, N., 2004. Assessing dimensions of perceived visual aesthetics of web sites. *International journal of human-computer studies*, 60(3), pp.269–298.

Mastin, L., 2008. The Basics of Philosophy. , pp.1–2. Available at: http://www.philosophybasics.com/branch_aesthetics.html [Accessed September 23, 2017].

Oyibo, K., Ali, Y.S. & Vassileva, J., 2016a. An empirical analysis of the perception of mobile website interfaces and the influence of culture. In *CEUR Workshop Proceedings*. pp. 44–56.

Oyibo, K., Ali, Y.S. & Vassileva, J., 2016b. Gender Difference in the Credibility Perception of Mobile Websites: A Mixed Method Approach. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization - UMAP '16*. Halifax, Canada, pp. 75–84. Available at: http://dl.acm.org/citation.cfm?doid=2930238.2930245.

Oyibo, K. & Hamada, M., 2013. A Framework for Instantiating Native Mobile Multimedia Learning Applications on Android Platform. *Proceedings of International Conference on Advances in Mobile Computing & Multimedia - MoMM '13*, pp.589–593. Available at: http://dl.acm.org/citation.cfm?doid=2536853.2536932.

Oyibo, K. & Vassileva, J., 2017. The Interplay of Aesthetics, Usability and Credibility in Mobile Website Design and the Moderation Effect of Gender. *SBC Journal on Interactive Systems*, 8(2), pp.4–19.

Oyibo, K. & Vassileva, J., 2016. The Interplay of Aesthetics, Usability and Credibility in Mobile Websites and the Moderation by Culture. In *Proceedings of the 15th Brazilian Symposium on Human Factors in Computer Systems - IHC '16*. pp. 1–10. Available at: http://dl.acm.org/citation.cfm?doid=3033701.3033711.

Perdue, R.R., 2001. Internet Site Evaluations: The Influence of Behavioral Experience, Existing Images, and Selected Website Characteristics. *Journal of Travel & Tourism Marketing*, 11(2–3), pp.83–103. Available at: http://www.tandfonline.com/doi/abs/10.1300/J073v11n02_05.

Pohlert, T., 2014. The pairwise multiple comparison of mean ranks package (PMCMR). R package. *R package*, pp.1–9. Available at: http://cran.r-project.org/package=PMCMR.

van Schaik, P. & Ling, J., 2009. The role of context in perceptions of the aesthetics of web pages over time. *International Journal of Human-Computer Studies*, 67(1), pp.79–89.

Sun, H., 2001. Building A Culturally-Competent Corporate Web Site : An Exploratory Study of Cultural Markers In Multilingual Web Design. *Language*, 10/01, pp.95–102. Available at: http://portal.acm.org/citation.cfm?doid=501516.501536.

Tractinsky, N., 2002. Visual Aesthetics. *Interaction-Design*, pp.1–71.

Wobbrock, J.O. et al., 2011. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 143–146.