

The Interplay of Aesthetics, Usability and Credibility in Mobile Website Design and the Effect of Gender

Kiemute Oyibo and Julita Vassileva
 Computer Science
 University of Saskatchewan
 Saskatoon, S7N 5C9, Canada
 kiemute.oyibo@usask.ca, jiv@cs.usask.ca

Abstract—In human-computer interaction, *aesthetics*, *usability* and *credibility* are key factors in the design of a successful website. Specifically, *aesthetics* has been identified as one of the main drivers of *web credibility*. However, in the mobile domain, large-scale research, cutting across cultures and continents, which is key to the generalizability of findings, is scarce. To bridge this gap, we conducted a multicultural study among 526 participants from 5 continents: Africa, Asia, North America, South America and Europe. Using four systematically designed mobile websites, we investigated: (1) the interrelationships among *aesthetics*, *usability* and *credibility*; and (2) the moderating effect of gender. Our results, based on partial least square path modeling, reveal that: (1) *perceived aesthetics* is stronger than *perceived usability* in predicting the *perceived credibility* of mobile websites; and (2) gender moderates the effect of *perceived aesthetics* on *perceived usability*, with this effect being stronger for males than for females. Our findings underscore the need for designers to pay closer attention to *aesthetics* in designing successful mobile websites, as their *visual appeal*, irrespective of gender, enhances their *perceived ease of use* and *credibility*. These findings are noteworthy because, given the usability challenges posed by the relatively small-screen size of the mobile device, designers may be tempted to focus on designing easy-to-use websites only, while downplaying their *visual appeal*. Such a decision may adversely impact the *overall credibility* of their websites going by users' first impression.

Keywords—*mobile website; path model; aesthetics; usability; credibility; gender; user interface, design layout, perception*

I. INTRODUCTION

The *perceived credibility* of a website is critical to the commercial success of its owners [1]. According to Fogg [2], “those who can design for credibility gain a strategic advantage” (p. 10) in the marketplace. *Web credibility* has been identified as an important factor in web design since it is the first point of contact with a business for new customers. As such, it is one of the most valuable assets of an organization. According to Fogg [2], *credibility* gives designers the power to change both the attitudes and behaviors of users. The high *credibility* of a website can foster positive user attitudes, such as thinking favorably of the website owners, feeling comfortable interacting with the website and embracing the point of view of the website owners. Similarly, a site's *credibility* can influence users' behaviors, for example, registering with the website by providing their personal

information, completing purchases of products and services and returning to the site some other day [3]. Moreover, research [4]–[7] has shown that *aesthetics* and *usability* are key factors in determining the *perceived credibility* of and *intention to use* websites. While a lot of research has been carried out on the influence of *aesthetics* and of *usability* [8] on *credibility* [1], [9], it remains to be seen in the mobile domain how these influences play out with respect to a mixed population comprised of different cultures, and how gender moderates the interrelationships among these three web design constructs [10]. Investigating this research topic has become important given: (1) the debate [11]–[16] regarding which of *aesthetics* (a hedonic attribute) and *usability* (an instrumental attribute) is more important, and thus should be given priority, in human-computer interaction (HCI) and user interface (UI) design; and (2) the temptation for designers to focus more on *usability* and less on *aesthetics* owing to the usability challenges the relatively small-screen size of the mobile device poses in human-computer interaction [17], [18].

In HCI, *aesthetics* and *usability* can be described as two key factors among a broad range of factors that influence the user experience. Morville [19] identified seven factors, which make up the User Experience Honeycomb. They include *usefulness*, *usability*, *findability*, *credibility*, *desirability*, *accessibility* and *value*. Of these factors, *desirability* (which relates to *aesthetics*), *usability* and *credibility* are among the most researched design constructs in the literature [1], [8], [9], [15], [20]–[23]. In particular, there have been many debates [12], [13], [22], [24], [25] between usability experts and graphics designers regarding which of *aesthetics* and *usability* is more important to users in judging a website based on their initial experience [26]. While some researchers and practitioners [24], [27] have argued that *usability* (a utilitarian factor) is very important in the design of HCI artifacts, others [1], [9], [22], [28] have stressed the need to pay attention to *aesthetics* (a hedonic factor) as well, as it is the first port of call in judging a website and the predominant determinant of *web credibility*. As argued by the Interaction Design Foundation [19], it is true that user experience as a discipline started with *usability*; however, it has outgrown *usability* to include other important factors, which designers must accommodate in order to deliver more effective and successful applications to their target users. Consequently, *aesthetics* has been identified as one of these factors, which is an important aspect of the *desirability* of a product, in general [19].

Furthermore, *aesthetics* has been identified as a key factor in the formation of users' first impression about a website, which is made within the first few seconds of encountering a website [29]. First impression is regarded as a key criterion in users' initial judgment of a website because it has to do with feelings and opinions quickly formed unconsciously on the basis of nothing other than *perceived aesthetics* [30]. According to Reinecke et al. [29], first impression is influential to the extent that it can later affect users' "*opinions of a site's usability and trustworthiness*" (p. 2049). According to Lidwell, Holden and Butler [31], "*aesthetic designs look easier to use and have probability of being used*" (p. 18). In general, aesthetic designs have been found to be more effective than less-aesthetic designs (which actually may be more usable) in fostering favorable user attitudes, including more tolerance of design problems [31]. Finally, research [32]–[34] has shown that gender differences exist in the evaluation of websites. Females are generally referred to as comprehensive information processors that pay attention to details, while males are referred to as selective information processors that miss subtle cues [35]. In HCI, females have been found to be driven by intrinsic factors more than males, while males by extrinsic factors more than females [36], [37].

However, most of the above findings regarding *aesthetics*, *usability* and *credibility*, on one hand, and the influence of gender, on the other hand, have been based on: (1) homogeneous samples; and (2) mostly studies (e.g., [10], [38]–[40]) conducted among the Western and Asian populations. These may threaten generalizability to a multicultural population sample. Moreover, there is limited research on how males and females differ [10], especially with respect to a multicultural population sample. To bridge these gaps, we conducted a large-scale empirical study among 526 subjects from five continents (Africa, Asia, North America, South America and Europe) to investigate the interrelationships among the three web design constructs and the influence of gender, using four systematically designed mobile websites as a case study. The results of our path modeling reveals that *perceived aesthetics* is stronger than *perceived usability* in determining the *perceived credibility* of mobile websites. Our results also reveal that there is a stronger aesthetic-usability effect for males than for females, indicating that males are more affected by the halo-effect phenomenon: a psychological cognitive bias in which the perception of one attribute of an object distorts the perception of other attributes [41], [42].

The rest of the paper is organized as follows. Sections II and III focus on the background and related work, respectively. Section IV and V focus on the research method and results, respectively. Finally, Sections VI, VII and VIII dwell on the discussion, limitation/future work and conclusion, respectively.

II. BACKGROUND

In this section, we provide an overview on the three web design constructs investigated in this paper.

A. Credibility

Credibility is defined as *believability* [4]. According to Fogg [43], there are four types of *web credibility*: *presumed*,

surface, *reputed*, and *earned credibility*. On one hand, *presumed credibility* is based on users' general assumptions, while *surface credibility* is based on their first impressions. On the other hand, *reputed credibility* is based on third-party endorsements/referrals, while *earned credibility* is based on a users' firsthand experience, consistent over time [43]. In this paper, we are concerned about *presumed/surface credibility*, which we refer to as *perceived credibility*. According to Fogg et al. [4], the *perceived credibility* of a website is composed of two main dimensions: the *perceived trustworthiness* of the owners and the *perceived expertise* of the designers. Robins and Holmes [1] observed that the judgment of the *credibility* of a website and its content places a burden on the user, since, unlike most journal publications, website information is not subjected to peer reviews or editorial processes. Hence, users adopt visceral means to base their assessment of *credibility*. Fogg et al. [4], postulated that, in evaluating the *credibility* of a website, users base their judgment on the assessment of *perceived trustworthiness* and *perceived expertise*. They explained that *credibility* is perceived by the users and does not reside in the evaluated object. In other words, *credibility* depends on the judgment users make about the object under evaluation. According to the prominence-interpretation theory, postulated by Fogg [44], users assess online *credibility* based on specific elements they notice in the object under evaluation (*prominence* of design elements) and the judgment they make about the observed elements (the *interpretation* they gave them) of the object. He went further to explain that, if a specific element in a website is not noticed by users, it will have no impact on their overall *credibility* judgment of the site. He identified five factors which may affect the *prominence* of a website element: *user involvement* with the website, *topic of the website*, *task of the user*, *experience of the user* and *individual differences* (ranging from the *need for cognition* to *level of literacy*). On the other hand, he identified *assumptions* in the user's mind (determined by culture, past experience, heuristics, etc.); *skill/knowledge* (determined by the user's level of competency on the website's subject matter); and *context* (e.g., the user's environment, expectations, situational norms, etc.) as the main factors that may influence the *interpretation* of the elements of a website under evaluation [44].

B. Aesthetics

Aesthetics is related to the notion of beauty. Historically, beauty is regarded as one of the ultimate values in Western philosophy, just like goodness, truth and justice. The nature of beauty has been one of the most fundamental and controversial subjects in philosophical aesthetics, which has endured several ages. It is a central theme among ancient Greek, Hellenistic and medieval philosophers as well as eighteenth and nineteenth century philosophers, such as Shaftesbury, Hutcheson, Kant, Hume and Hegel. Although, in the beginning of the twentieth century, the study of beauty as a subject of philosophical inquiry and as a primary goal of the art witnessed a decline, a new wave of interest began gaining traction in the early years of the twenty-first century [45]. In particular, Baumgarten reappraised the concept of *aesthetics*, which, prior to the mid-eighteenth century, had meant "sensibility" ("*responsiveness to the stimulation of the senses*") [46], [47]. He redefined it as the judgment of taste or that which is beautiful. He based the

judgment of taste or beauty on the feeling of pleasure or displeasure. Hence, *aesthetics* is defined as the branch of philosophy, which is concerned with “*the nature and appreciation of art, beauty and good taste*” [48]. It is derived from the Greek word “*aisthetikos*”, which means “*of sense perception*” [48]. Moreover, it is part of axiology, which is the study of value judgment. According to [48], the judgment of aesthetic value relies on our ability to discriminate at the sensory, emotional and intellectual levels. However, while Baumgarten referred to the judgment of beauty as cognitive, which has to do with rational ideas, such as harmony, Kant referred to it as affective, which has to do with the form of the object under evaluation [49], [50]. Consequently, Kant emphasizes that “*aesthetics judgment must concern itself only with form (shape, arrangement, rhythm, etc.) in the object presented, not with sensible content (color, tone, etc.)*” [49].

Furthermore, according to Kant, the *aesthetic judgment* of an object is a subjective experience, which is *disinterested*, *universal* and *necessary*. It is disinterested in the sense that it is desire-free, meaning we feel or take pleasure in an object because we judge it beautiful, and not the other way round, as held by Baumgarten. On the other hand, it is universal and necessary because the beauty of an object is self-evident and demands others agree with us on its being beautiful. However, according to Kant, although beauty appears as if it were an actual property of the object, it possesses no objective property that makes the object under evaluation beautiful. Rather, the universality and necessity (self-evidence) of aesthetic judgment are a product of the features of the human mind, which Kant called “*common sense*” [49], [51].

In the context of HCI design, *aesthetics* is defined as the visual or aesthetic appeal of an artifact (e.g., a website) [52]. Though, in general, there are many forms of *aesthetics* (e.g., “*literary aesthetics*”), in this paper, we are particularly concerned with “*visual aesthetics*” as it relates to web design and the visual sense of perception of users. Thus, we also refer to *visual aesthetics* as “*visual appeal*” and/or “*perceived aesthetics*” from the users’ point of view, just like *perceived usability* and *perceived credibility*. Though, historically, *aesthetics* is associated with order and form, in more recent literature, it is associated with delight and perception [21]. More specifically, *perceived aesthetics* in HCI research is related to the notion of *aesthetic judgment*, which we covered in the foregoing paragraph.

Based on Lavie and Tractinsky’s [21] psychometric analyses on developing a scale for HCI research, it was found that *perceived aesthetics* is composed of two dimensions: *classical* and *expressive aesthetics*. *Classical aesthetics* pertains to the traditional notion of orderliness and cleanness of the design of an artifact, while *expressive aesthetics* reflects the originality and creative power of the designer of the artifact. In HCI and web design, *classical aesthetics* is operationalized by the use of simple terms such as “*visual*,” “*pleasant*,” “*clean*,” etc., while *expressive aesthetics* is operationalized by the use of complex terms such as “*fascinating*,” “*creative*,” “*sophisticated*,” etc. [21], [53], [54]. Furthermore, *perceived aesthetics* may depend on a number of factors, such as gender, culture, etc. Thus, for example, what one culture/gender may

perceive as aesthetic may be perceived as unaesthetic by another culture/gender [42].

C. Usability

Usability is defined as the *ease of use* of an information system [55]. In other words, it captures how effective and free of efforts humans can interact with an information system (e.g., website) [56]. Basically, we have two types of *usability*: *perceived* and *actual usability* [13]. Empirically, *perceived usability* can be measured by asking users to give their opinions about their perceived ease of use of a system. Thus, it can be operationalized using terms such as “*easy to use*”, “*easy to navigate*”, “*convenient to use*”, etc. [21]. On the other hand, *actual usability* can be measured objectively by recording how well users can master and perform tasks using a system [56].

III. RELATED WORK

A number of studies, empirical and experimental, have been carried out to investigate the interrelationships among *aesthetics*, *usability* and *credibility* (“*perceived*” omitted from the names of the constructs, for brevity, here and other parts of the paper). In this section, we review related works in this area.

A. Aesthetics-Usability Relationship

There have been a lot of studies on the relationship between *aesthetics* and *usability* (aka “*aesthetic-usability effect*” [31]). Kurosu and Kashimura [28] conducted an experimental study among 252 students in Japan to investigate the determinants of the *perceived usability*, which they called *apparent usability*, of an information system. The study was based on an Automated Teller Machine (ATM), which had 26 different layouts. Their correlation analysis revealed that the *apparent usability* of the ATM was more strongly impacted by its *perceived aesthetics* than by its *inherent usability*. Thus, they concluded that *perceived aesthetics* influences *perceived usability* more than *inherent usability* does. As a follow-up, Tractinsky [22] conducted a similar study among 104 first-year engineering students in Israel, using the same study materials (26 ATM layouts translated into Hebrew from Japanese). The study was carried out in an attempt to replicate Kurosu and Kashimura’s [28] study and uncover the effect of culture. Tractinsky [22], just like Kurosu and Kashimura [28], found a strong correlation between *perceived aesthetics* and *perceived usability*, suggesting that the aesthetic-usability effect holds, irrespective of culture. Other studies (e.g., [13], [15], [23], [57]) have also confirmed the relationship between *aesthetics* and *usability*. For example, in a further follow-up experimental study, Tractinsky, Katz and Ikar [13] investigated the determinants of the *perceived usability* of nine of the 26 adapted ATM layouts after users had actually interacted with them. They found that the *perceived usability* of the ATM after use was influenced mainly by its *perceived aesthetics* before use and not by its *actual usability*. Thus, they concluded “*what is beautiful is usable*.” They attributed this finding to the “*halo effect*” [41], [42]. However, Hassenzahl [58], who questioned this finding, upon attempting to confirm it in a similar experimental study, reported that he was unable to replicate the halo-effect phenomenon. Further, in the mobile domain, Li and Yeh [59], carried out a study on *mobile trust* using e-commerce websites

and 200 Taiwanese subjects as a case study. Just as in earlier research [13], [28] in the non-mobile domain, they found that *design aesthetics* significantly impacts *perceived ease of use*. Finally, in mobile website design, Oyibo and Vassileva [23] investigated the relationship between *classical aesthetics* vs. *expressive aesthetics* and *perceived usability*. They found that, regardless of culture, *classical aesthetics* has a stronger influence than *expressive aesthetics* on *perceived usability*.

B. Aesthetics-Credibility Relationship

A number of studies have also been carried out on the link between *aesthetics* and *credibility*. Robins and Holmes [1] conducted a study of 42 websites among 20 Information Science students from Kent State University in the United States. They found that when the same information content is presented to users at different levels of aesthetic treatment, the contents with better *visual presentation* is likely to be perceived as having higher *credibility*. In the mobile domain, Oyibo, Ali and Vassileva [60] replicated this finding in a mixed-method study of mobile websites with different levels of *aesthetic appeal* among 285 participants from Africa, North America and Asia. Further, Fogg et al. [9] evaluated the *credibility* of two actual (health) websites in a qualitative study among over 2500 subjects from around the world, 71% of whom were from the United States. They found that *design look* (which is related to *perceived aesthetics*) is the primary determinant of the *perceived credibility* of the investigated websites. Similarly, Liu, Lee and Lee [6] conducted a study among 65 junior and sophomore students from the Department of Digital Media Design in a Taiwanese university on the connection between *perceived aesthetics* (which they referred to as *appearance*) and *perceived credibility*. Their study was based on the empirical evaluation of a number of online banking websites. They found that better-looking websites were rated as more credible than less-good-looking ones. Further, in their investigation of mobile websites using partial least square path modeling and Canada and Nigeria as a case study, Oyibo and Vassileva [7] found a strong relationship between *perceived aesthetics* and *perceived credibility*.

C. Usability-Credibility Relationship

A number of studies have investigated the relationship between *usability* and *credibility* in web design. Fogg et al. [4] conducted a study of 51 websites among over 1400 participants from the United States and Europe to determine the factors that drive the *perceived credibility* of websites. They found *ease of use* (coming behind *real world feel*) as the second strongest determinant of *perceived credibility*. Li and Yeh [61] also conducted a study among 200 Taiwanese on mobile commerce websites. They found a significant relationship between *perceived ease of use* and *mobile trust*, which is a component of *mobile credibility* [4]. Youngblood and Mackiewicz [16] conducted a usability evaluation of government websites in the United States. They found that *perceived usability* impacts the *credibility* of the investigated websites. Similarly, Oyibo and Vassileva [62] conducted a study among 233 participants from Canada and Nigeria on mobile *web credibility*. They found a significant relationship between *perceived usability* and *perceived credibility*, which cuts across culture. However, in Neumark et al.'s [63] investigation on the quality and impact of

online health information on users, using 29 Hebrew websites as a case study, they found no correlation between the *perceived usability* and the *perceived credibility* of the investigated websites.

D. Effect of Gender in Web Design and Credibility

Limited research has been done on the impact of gender on the judgment of websites in the mobile domain. In the desktop domain, Cyr [10] conducted a study on website design and gender difference across eight different cultures (Canada, United States, India, Germany, Japan, Mexico, Chile, and China) among 1156 participants. Similarly, in the mobile domain, Oyibo and Vassileva [60] conducted a study among 285 participants from North America, Africa and Asia. While both groups of authors found that females are more critical than males in evaluating websites, they did not investigate how gender moderates the interrelationships among the three web design constructs: *aesthetics*, *usability* and *credibility*, which we are currently investigating in this paper. In fact, most prior studies (e.g., [32], [33], [64]–[66]) in the literature often focus on how both genders differ in terms of participants' preferences and the levels of perception of certain design attributes—and neither in terms of the interrelationships among the three web design constructs nor the influence of one design construct on another in a partial least square path model or structural equation model, where constructs compete against one another in terms of magnitude of influence or effect size. Moreover, most prior studies in the extant literature have been:

1. concentrated on subjects mainly from Western and Asian cultures (e.g., [6], [9], [10], [28]);
2. based on monocultural and convenience samples, involving mainly student subjects (e.g., [1], [6], [22], [28]); and
3. focused on the desktop domain (e.g., [1], [16], [64], [66]).

These limitations may threaten the generalizability of findings to the mobile domain and a mixed population sample, which cuts across different cultures and geographical locations. Our study attempts to bridge these gaps by using a more globally diverse population sample, which includes subjects from different cultures and continents, including understudied continents such as Africa and South America.

IV. METHOD

This section covers our research question, research design, measurement instruments and participants' demographics.

A. Research Questions

In this paper on the perception and evaluation of mobile web design, using a multicultural population sample, we aim to answer the following two main research questions:

1. *How do the three important web design constructs (aesthetics, usability and credibility) influence one another in the mobile web design?*
2. *Do the interrelationships among aesthetics, usability and credibility depend on gender and/or UI design?*

B. Research Design

To answer our research questions, we systematically designed four versions of a hypothetical mobile website¹ based on UI designs adapted from the marketplace (see Fig. 1 [67]). The websites comprise two basic types of designs: low-aesthetic design websites (WA and WB) and high-aesthetic design websites (WC and WD). We arrived at these four websites using a systematic transformation framework we called artifact-action framework. The quadrants of the framework represent the artifacts (web designs), while the axes represent the actions (manipulations) carried out on one web design (in a given quadrant) to arrive, in a clockwise fashion, at a new web design (in the next quadrant). For example, starting from WA (a multi-color design), to arrive at WB (a minimalistic design), we carry out the action “make gray and add icon.” Further, to arrive at WC (a blue-theme design) from WB, we carry out the action “make unicolor.” This continues in a clockwise manner until we arrive at the initial web design (WA) we started from.

C. Research Hypotheses

Based on the path model in Fig. 2, we formulated eight hypotheses, relating to the four UIs in Fig. 1, as follows:

H1: The direct effect of *Aesthetics* (*A*) on *Credibility* (*C*) is stronger than the direct effect of *Usability* (*U*) on *Credibility* (*C*).

H2: The direct effect of *Aesthetics* on *Usability* is stronger for the female group than for the male group.

H3: The direct effect of *Aesthetics* on *Credibility* is stronger for the female group than for the male group.

H4: The effect size of *Aesthetics* on *Credibility* is stronger for the female group than for the male group.

H5: The direct effect of *Usability* on *Credibility* is higher for the male group than for the female group.

H6: The effect size of *Usability* on *Credibility* is higher for the male group than for the female group.

H7: The mediation of the direct effect of *Aesthetics* on *Credibility* by *Usability* is higher for the grid-layout mobile design than the list-layout mobile designs.

H8: The mediation of the direct effect of *Aesthetics* on *Credibility* by *Usability* is stronger for the male group than for the female group.

We used a within-subject design, with all four mobile webpages (UIs) presented to each participant in this order—WC, WA, WB and WD—in the online survey. To minimize the possible impact of the presentation order, we specifically ensured that the mobile webpages were not presented (based on

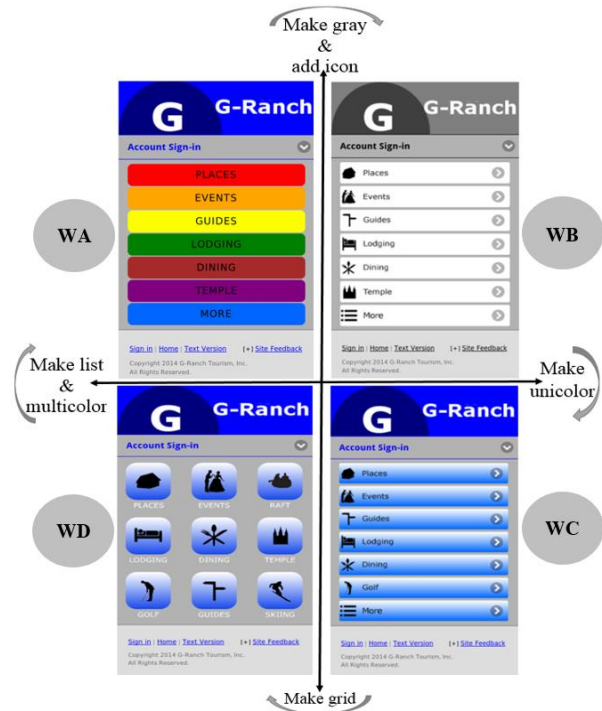


Fig. 1. Systematically designed user interfaces

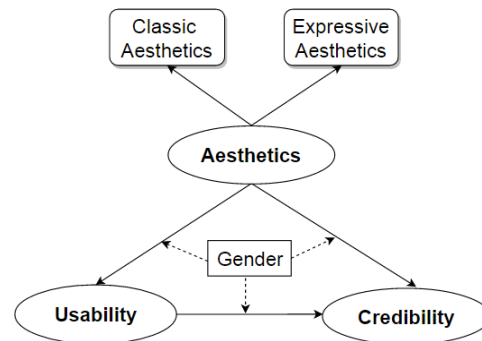


Fig. 2. Hypothesized path model.

our perception and evaluation) from the most aesthetic to the least aesthetic or vice versa, e.g., starting with WD and WC and ending with WB and WA, or vice versa.

The hypothetical model in Fig. 2, as well as the eight hypotheses above, are informed by prior findings in the literature. The first hypothesis (H1) was informed by the findings by Robins and Holmes [1], Fogg et al. [9] and other researchers [7], [15] in the desktop domain, with respect to *web design* and *credibility*. The authors, in different studies, found that *perceived aesthetics* is the predominant design attribute that determines the *credibility* of a website. For example, according to Robins and Holmes [1], *aesthetics* is the main factor that may influence a user to either stay on a website upon the first visit or move to another. The second to the sixth hypotheses (H2 to H6) were based on the assumption

¹We adapted all four websites from existing tourism and travels' mobile websites (m.wakanow.com, mobile.united.com, mobile.utah.com and tourismwinnipeg.com) in the market in 2014. However, some of the websites' UIs must have been redesigned [72].

that if a given gender is more concerned about any of the two exogenous constructs (*aesthetics* or *usability*) than the other gender, then the influence of that exogenous construct on the endogenous constructs (*usability* or *credibility*) will be stronger for that gender than the other. Generally, research has shown that men are more motivated by extrinsic factors (e.g. *usability*), which are task-oriented, than women, while women, on the other hand, are more motivated by intrinsic factors, which are hedonic in nature (e.g., *aesthetics*), than men [68]. Thus, with respect to H2, H3 and H4, we hypothesized that the influence and effect size of *aesthetics* on *credibility* (H3 and H4, respectively), on one hand, and the influence of *aesthetics* on *usability* (H2), on the other hand, will be stronger for females than for males because prior research [7], [69], [70] has shown that females care about *aesthetics* more than males do. As such, we hypothesized that *aesthetics* may have more effect on the judgment of women than men [71]. With respect to H5 and H6, we hypothesized that the influence and effect size of *usability* on *credibility* will be stronger for males than females because prior research (e.g., [60], [69]) has shown that males care about *usability* more than females.

Furthermore, the seventh hypothesis (H7) was based on the fact that mobile WD possesses a unique layout, which is different from the other UIs. Thus, we hypothesized that the grid layout (having more spaced-out items, which makes it more usable or less prone to click-errors than the others [72]) will result in *perceived usability* influencing *perceived credibility* more for mobile WD than for the other mobile webpages, which has a list layout (having less spaced-out items, which may cause more click-errors). Thus, as stated in the seventh hypothesis (H7), we expected to have *usability* mediating the influence of *aesthetics* on *credibility* more for mobile WD than for mobile WA, WB and WC. Finally, the eighth hypothesis (H8) is formulated based on the same assumption made regarding the fifth hypothesis (H5). Basically, our eighth hypothesis (H8) is that, if the influence of *usability* on *credibility* is stronger for the male group than the female group, then, *usability* will mediate the influence of

aesthetics on *credibility* for the male group more than for the female group.

D. Measurement Instruments

To measure all three web design constructs in our hypothesized path model, we used existing scales that have been previously validated. To measure *perceived aesthetics*, we used the six-item scale of Lavie and Tractinsky [5] as adapted by van Schaik and Ling [73]. The *aesthetics* scale comprises two subscales: *classical aesthetics* and *expressive aesthetics*. Similarly, to measure *perceived usability*, we used Lavie and Tractinsky's five-item scale [5] as adapted by De Angeli, Sutcliffe and Hartmann [74]. Finally, to measure *perceived credibility*, we used a single-item scale [75], which has been shown by prior research [76], [77] to be as reliable as a multi-item scale. Each item in each of the scales was measured using a 7-point Likert scale, ranging from "Strongly Disagree (1)" to "Strongly Agree (7)." In administering the online survey, all of the items from both *aesthetics* and *usability* scales for each mobile webpage were presented together in a randomized fashion to each participant. This was done with the intention to prevent participants from easily identifying the design construct we are measuring at a given instance in time, e.g., if we presented each construct's items separately in different blocks under the image of each webpage [23]. The *credibility* question was asked immediately after the randomized questions on *aesthetics* and *usability* had been asked.

E. Participants

Our study was submitted to our university's research ethics board for evaluation and approval. Thereafter, upon approval, we invited and recruited participants, using our university's website, Facebook and email, to take part anonymously in our online survey. In order to appreciate them for their time, participants were given an opportunity to optionally enter for a draw for a chance to win one of our four gift cards of C\$50 each. About 550 participants took part in the study as a whole. However, after cleaning the data, we were left with a total of 526 participants for our final analysis. Table I shows the demographic information of the participants, which comprised 53.6% males and 44.5% females.

TABLE I. PARTICIPANTS' DEMOGRAPHICS ACCORDING TO COUNTRIES.

Criterion	Group	Nigeria	Ghana	Brazil	China	Canada	Others	Overall	Perc.
Gender	Male	115	37	34	46	25	25	282	53.6%
	Female	40	13	16	63	71	31	234	44.5%
	Unidentified	1	0	0	6	2	1	10	1.9%
Age	18-24	118	6	16	51	55	18	264	50.2%
	25-34	32	40	21	59	31	31	214	40.7%
	>34	2	3	13	5	12	6	41	7.8%
	Unidentified	4	1	0	0	0	2	7	1.3%
Years on Internet	<10	116	12	8	82	13	9	230	43.7%
	>=10	50	38	42	33	85	48	296	56.3%
Educational Qualification	Technical/Trade	8	0	1	12	4	1	26	4.9%
	High School	108	2	8	9	40	10	177	33.7%
	Bachelor	21	36	13	67	42	13	192	36.5%
	Postgraduate	11	9	28	25	11	31	115	21.9%
	Unidentified	8	3	0	2	1	2	16	3.0%
	Subtotal	156	50	50	115	98	57	526	100%
	National Percent	29.7%	9.5%	9.5%	21.9%	18.6%	10.8%	100%	

V. RESULT

In this section, we present the results from our Partial Least Square Path Modeling (PLSPM) using R's *plspm* package [78].

A. Evaluation of Global Measurement Model

One of the requirements for the analysis of the structural model is the evaluation of the measurement (inner) model [78].

For the global model, we present the evaluation of the measurement model for all three constructs under investigation (*aesthetics*, *usability* and *credibility*) as shown in Table II.

1) *Indicator Reliability*: Most of the indicators in the measurement models have an outer loading greater than 0.7 [6]. As such, the reliability criterion is met, as the communality values for most indicators are greater than 0.5.

TABLE II. INTERNAL CONSISTENCY, INDICATOR RELIABILITY, AND CONSTRUCT VALIDITY.

Webpage	Block	Indicator	Loading	Communality	AVE	DG.rho
WA	Aesthetics	Classical lower order indicator	0.96	0.92	0.91	0.95
		Expressive lower order indicator	0.95	0.91		
	Classical Aesthetics	Clean	0.89	0.80	0.81	0.93
		Pleasant	0.91	0.83		
		Visual	0.89	0.79		
	Expressive Aesthetics	Creative	0.89	0.79	0.75	0.90
		Fascinating	0.91	0.82		
		Sophisticated	0.80	0.64		
	Usability	Clear design	0.81	0.66	0.72	0.93
		Convenient to use	0.88	0.78		
		Easy to navigate	0.87	0.77		
		Easy orientation	0.84	0.71		
	Credibility	Easy to use	0.84	0.71	1.00	1.00
		Credibility level	1.00	1.00		
WB	Aesthetics	Classical lower order indicator	0.95	0.90	0.89	0.94
		Expressive lower order indicator	0.94	0.88		
	Classical Aesthetics	Clean	0.78	0.61	0.74	0.89
		Pleasant	0.90	0.81		
		Visual	0.89	0.80		
	Expressive Aesthetics	Creative	0.89	0.80	0.71	0.88
		Fascinating	0.89	0.80		
		Sophisticated	0.73	0.54		
	Usability	Clear design	0.81	0.65	0.74	0.94
		Convenient to use	0.87	0.76		
		Easy to navigate	0.88	0.77		
		Easy orientation	0.86	0.74		
	Credibility	Easy to use	0.88	0.78	1.00	1.00
		Credibility level	1.00	1.00		
WC	Aesthetics	Classical lower order indicator	0.96	0.91	0.90	0.95
		Expressive lower order indicator	0.94	0.88		
	Classical Aesthetics	Clean	0.79	0.62	0.75	0.90
		Pleasant	0.90	0.82		
		Visual	0.91	0.82		
	Expressive Aesthetics	Creative	0.90	0.82	0.70	0.88
		Fascinating	0.90	0.82		
		Sophisticated	0.69	0.47		
	Usability	Clear design	0.84	0.71	0.68	0.92
		Convenient to use	0.85	0.72		
		Easy to navigate	0.81	0.65		
		Easy orientation	0.82	0.67		
	Credibility	Easy to use	0.82	0.67	1.00	1.00
		Credibility level	1.00	1.00		
WD	Aesthetics	Classical lower order indicator	0.96	0.92	0.90	0.95
		Expressive lower order indicator	0.94	0.88		
	Classical Aesthetics	Clean	0.88	0.77	0.81	0.93
		Pleasant	0.92	0.84		
		Visual	0.91	0.83		
	Expressive Aesthetics	Creative	0.89	0.79	0.70	0.87
		Fascinating	0.88	0.77		
		Sophisticated	0.72	0.52		
	Usability	Clear design	0.87	0.76	0.79	0.95
		Convenient to use	0.88	0.77		
		Easy to navigate	0.89	0.79		
		Easy orientation	0.90	0.82		
	Credibility	Easy to use	0.91	0.83	1.00	1.00
		Credibility level	1.00	1.00		

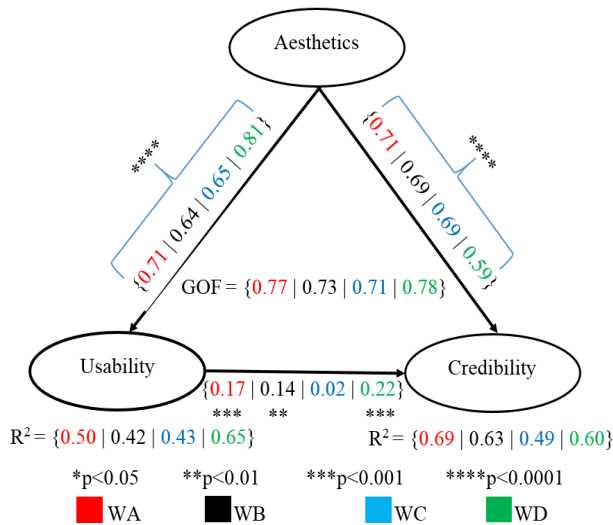


Fig. 3. Global path model.

2) *Internal Consistency Reliability*: The internal consistency reliability of the constructs was evaluated using the composite reliability criterion, DG.rho (ρ), which is greater than 0.7 for all of the constructs, as shown in the last column of Table II [79].

3) *Convergent Validity*: The Average Variance Extracted (AVE) was used to assess convergent validity. The AVE for each model construct was above the recommended value, i.e., greater than 0.5 [79].

4) *Discriminant Validity*: The crossloading of each construct was also assessed to determine the discriminant validity of each construct. No indicator loaded higher on any other construct than the one it was meant to measure [79].

B. Evaluation of Subgroup Measurement Models

Our multigroup analysis (MGA) based on gender shows that there is a statistically significant difference between males and females with respect to specific relationships in the global model (see Table III). This led us to build specific models for the male and female subgroups. As such, just as we did for the global model, we evaluated the measurement models for both subgroups. The result showed that both inner models met the required criteria (assessed and presented in the previous subsection) as well. Due to limited space, and for the sake of brevity, the results are not included in the paper.

C. Analysis of Global Model

Upon establishing that the measurement models met the required assessment criteria, we carried out an evaluation of the structural models shown in Fig. 3 for all four mobile webpages {WA | WB | WC | WD}. The evaluation analysis of the structural models was based on the overall (global) data. For brevity, we have compacted all four global models for all four mobile webpages into one overarching model for quick and easy visualization and side-by-side comparison across all four models.

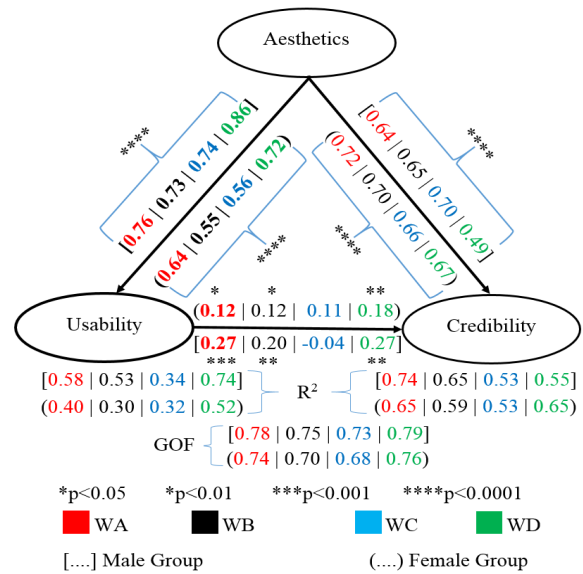


Fig. 4. Subgroup path model for male and female groups (bold fonts indicate paths where there are significant differences between males and females).

The path model shows: (1) the direct effects (β) exerted by the exogenous constructs (independent variables) on the endogenous constructs (dependent variables); (2) the coefficients of determination (R^2) of the endogenous constructs (defined as the amount of variance in an endogenous construct explained by the exogenous constructs); and (3) the goodness of fit (GOF), which indicates the predictive power of each model or how well each model fits its data [78]. The number of asterisks indicates how significant each of the direct effects is, with one, two, three and four asterisks representing the p-values of 0.05, 0.01, 0.001 and 0.0001, respectively. The p-value of 0.05 indicates our chosen minimum confidence level ($\alpha = 95\%$) for the direct effect, also known as influence or path coefficient (β) in our paper.

Fig. 3 reveals that the global models fit their respective data well due to the high GOF > 0.7 for all four models, with mobile WA and WD having the highest values of 0.77 and 0.78, respectively, and mobile WB and WC having the least values of 0.73 and 0.71, respectively. Further, based on the PLS-PM standard, $R^2 < 0.30$ is low, $0.30 < R^2 < 0.50$ is moderate and $R^2 > 0.60$ is high [78]. Thus, the R^2 for *usability* is moderate for mobile WA, WB and WC (50%, 42% and 43%, respectively) but high for mobile WD (65%). This indicates that the model accounts for more variance in *usability* for mobile WD (65%) than for the other mobile webpages. For *credibility*, the R^2 values range from 49% (moderate) for WC to 69% (high) for mobile WA, indicating more of the variance in *credibility* for mobile WA is accounted for than for other webpages. In summary, the global models have high predictive power and explain moderate-to-high amount of the variance in the endogenous constructs (*usability* and *credibility*). Further, we see that all the direct effects are significant, except for the direct effect of *usability* on *credibility* [(U→C)_{DIR}] for mobile WC, with $\beta = 0.02$. The path coefficients (β) are highly significant for (A→U)_{DIR} and (A→C)_{DIR} paths (ranging from

0.59 to 0.81 at $p < 0.0001$ for mobile WD) and quite significant for $(U \rightarrow C)_{DIR}$ paths (ranging from 0.14 at $p < 0.01$ for mobile WB to 0.22 at $p < 0.001$ for mobile WD). However, there was need for us to find out whether these path coefficients were moderated by gender by carrying out a multigroup analysis.

D. Multigroup Analysis

From the multigroup analysis (MGA) we conducted for the male and female subgroups, we found statistically significant differences in the $(A \rightarrow U)_{DIR}$ path for all four mobile webpages and in the $(U \rightarrow C)_{DIR}$ path for mobile WA only (at $p < 0.05$), as shown in the MGA results in Table III and indicated by the bold fonts in Fig. 4. In all five paths, where significant differences exist, the coefficients are stronger for the male group than for the female group. For example, for the $(A \rightarrow U)_{DIR}$ path, the male and female groups have path coefficients [0.76 | 0.73 | 0.74 | 0.86] and [0.64 | 0.55 | 0.56 | 0.72] at $p < 0.0001$ for mobile WA, WB, WC and WD, respectively. Similarly, for the $(U \rightarrow C)_{DIR}$ path, the male and female groups have path coefficients of 0.27 at $p < 0.001$ and 0.12 at $p < 0.05$, respectively, for mobile WA.

E. Direct, Indirect and Total Effects

As shown in Fig. 5, we plotted the direct effects $[(A \rightarrow C)_{DIR}]$ and $[(U \rightarrow C)_{DIR}]$, the indirect effects $[(A \rightarrow C)_{IND}]$ and the total effects $[(A \rightarrow C)_{TOT}]$ in order to easily visualize the trend of effects as the web design is systematically manipulated from the least aesthetic design (WA) to the most aesthetic design (WD). (See [67] for the aesthetic ratings.) A remarkable trend in the plot is that the total effect trace $[(A \rightarrow C)_{TOT}]$ for all four webpages is in sync with the direct effect $[(U \rightarrow C)_{DIR}]$ and indirect effect $[(A \rightarrow C)_{IND}]$ traces, with a sharp dip recorded at WC in the global and male models. This reflects the fact that, analytically, the total effect $[(A \rightarrow C)_{TOT}]$ is a function of the direct effect $[(A \rightarrow C)_{DIR}]$ and indirect effect $[(A \rightarrow C)_{IND}]$, while the indirect effect $[(A \rightarrow C)_{IND}]$ is a function of the direct effects $[(A \rightarrow U)_{DIR}]$ and $[(U \rightarrow C)_{DIR}]$. Another interesting finding common to all three plots (in Fig. 5) is that there is a significant indirect effect of *aesthetics* on *credibility* via *usability* for mobile WA, WB and WD, but none for mobile WC. This indirect effect $[(A \rightarrow C)_{IND}]$ is relatively stronger in the global and male models. We provide a possible explanation for this finding in the discussion of the eighth hypothesis (H8) in Section VI.

F. Mediation

We analyzed the global and subgroup models to determine whether *usability* mediates the effect of *aesthetics* on *credibility*. As shown in Fig. 6, for both models, we computed and plotted a measure of mediation, variance accounted for (VAF), which represents the ratio of the significant indirect effect to the total effect on *credibility* [79]. Based on the recommended guideline [79], which states that $VAF < 0.2$ indicates no mediation, $0.2 \leq VAF \leq 0.8$ indicates partial mediation and $VAF > 0.8$ indicates full mediation, we found that mediation by *usability* only occurred for: (1) the global model with respect to mobile WD, in which the VAF is 0.23; and (2) the male model with respect to mobile WA and WD, in which the VAFs are 0.24 and 0.32, respectively. It is

noteworthy that there is no mediation for: (1) mobile WC because the $(U \rightarrow C)_{DIR}$ path coefficient in Fig. 3 and Fig. 4 is non-significant; and (2) mobile WB because $VAF < 0.20$, as shown in Fig. 6.

TABLE III. MULTIGROUP ANALYSIS BETWEEN MALE AND FEMALE

Webpage	Path	Global	M	F	Sig
WA	$(A \rightarrow U)_{DIR}$	0.71	0.76	0.64	0.05
	$(A \rightarrow C)_{DIR}$	0.71	0.64	0.72	n.s
	$(U \rightarrow C)_{DIR}$	0.17	0.27	0.12	0.05
WB	$(A \rightarrow U)_{DIR}$	0.67	0.73	0.55	0.001
	$(A \rightarrow C)_{DIR}$	0.69	0.65	0.70	n.s
	$(U \rightarrow C)_{DIR}$	0.15	0.20	0.12	n.s
WC	$(A \rightarrow U)_{DIR}$	0.65	0.74	0.56	0.001
	$(A \rightarrow C)_{DIR}$	0.69	0.70	0.66	n.s
	$(U \rightarrow C)_{DIR}$	0.02	-0.04	0.11	n.s
WD	$(A \rightarrow U)_{DIR}$	0.81	0.86	0.72	0.001
	$(A \rightarrow C)_{DIR}$	0.59	0.49	0.67	n.s
	$(U \rightarrow C)_{DIR}$	0.22	0.27	0.18	n.s

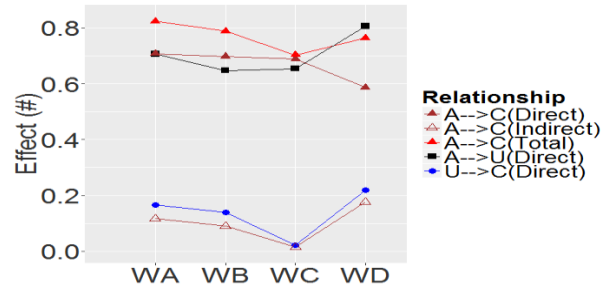


Fig. 5a. Direct, indirect and total effect for the global model

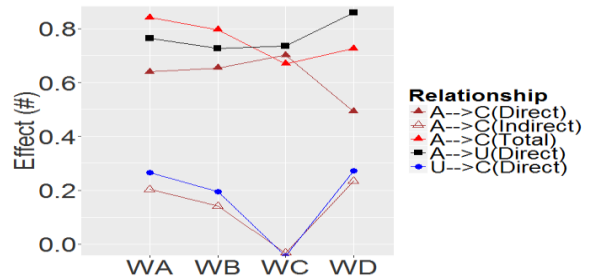


Fig. 5b. Direct, indirect and total effect for the male model

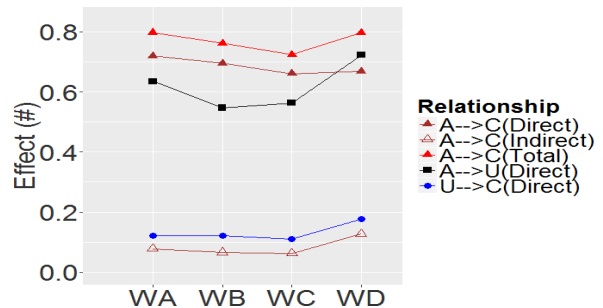


Fig. 5c. Direct, indirect and total effect for the female model

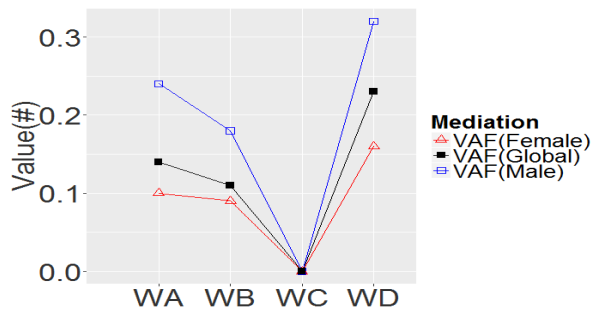


Fig. 6. Usability mediation effect of aesthetics on credibility

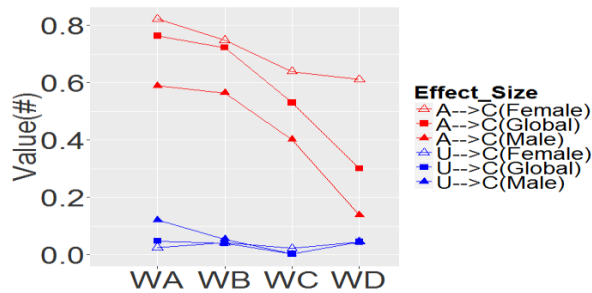


Fig. 7. Effect sizes of driver constructs on credibility

G. Effect Size

We computed the effect sizes of both exogenous/driver constructs (*aesthetics* and *usability*) on the endogenous/target construct (*credibility*). We used the formula recommended by Hair et al. [79]. According to Preacher and Kelly [80], it has become very important for researchers to report the effect sizes (f^2 or ES) of the driver constructs on the target construct whenever possible. Cohen [81] defined effect size as the extent to which the phenomenon of interest is present among the studied population. Effect size could also be defined as a measure of *how strong* the relationship between two variables of interest is [82]. Fig. 7 shows a line graph of the effect sizes for the driver constructs (*aesthetics* and *usability*) on *credibility*. According to Cohen’s guidelines [81], effect sizes of 0.02, 0.15 and 0.35 are classified as small, medium and large, respectively. As shown in the graph, the effect size of *aesthetics* on *credibility* [(A→C)_{ES}] is predominantly large, while that of *usability* on *credibility* [(U→C)_{ES}] is predominantly small. For example, with respect to WA | WB | WC | WD, the effect sizes [(A→C)_{ES}] for females are (0.82 | 0.74 | 0.64 | 0.61), respectively, while those for males are [0.59 | 0.56 | 0.40 | 0.14], respectively. On the other hand, the effect sizes [(U→C)_{ES}] for males are [0.12 | 0.05 | 0.002 | 0.04], respectively, while those for females are (0.02 | 0.05 | 0.02 | 0.05), respectively. Firstly, the predominantly all-high effect sizes [(A→C)_{ES}], except that with respect to mobile WD for the male group, compared to the predominantly all-small effect sizes [(U→C)_{ES}] for both genders, suggests that, irrespective of gender, *aesthetics* is more important than *usability* in predicting *credibility*. Secondly, with respect to (A→C)_{ES}, the effect size is higher for females than for males.

This suggests that females care about *aesthetics* in evaluating *credibility* more than males do.

H. Verification of Hypotheses

In this subsection, we present the verification of our hypotheses.

1) *The direct effect of Aesthetics on Credibility is stronger than the direct effect of Usability on Credibility:* As shown in Fig. 3 and Fig. 4, the direct effects of *aesthetics* on *credibility* are relatively higher than the direct effects of *usability* on *credibility* at both the global and subgroup levels. For example, at the global level, the direct effects for (A→C)_{DIR} {0.71 | 0.64 | 0.65 | 0.81} with respect to {WA | WB | WC | WD} are relatively higher than the direct effects for (U→C)_{DIR} {0.17 | 0.14 | 0.02 | 0.22}. Similarly, for the male group, the direct effects for (A→C)_{DIR} [0.76 | 0.73 | 0.74 | 0.86] are relatively higher than the direct effects for (U→C)_{DIR} [0.27 | 0.20 | -0.04 | 0.27]. Finally, for the female group, the direct effects for (A→C)_{DIR} (0.64 | 0.55 | 0.56 | 0.72) are relatively higher than the direct effects for (U→C)_{DIR} (0.12 | 0.12 | 0.11 | 0.18). Based on these results, we conclude that our first hypothesis (H1) is fully supported and replicated across both genders and all UI designs.

2) *The direct effect of Aesthetics on Usability is stronger for the female group than for the male group:* As shown in Fig. 4, contrary to our hypothesis, the influence of *aesthetics* on *usability* is higher for the male group [0.76 | 0.73 | 0.74 | 0.86] than for the female group (0.64 | 0.55 | 0.56 | 0.72). However, the difference between each corresponding pair of path coefficients is not significant (see Table III). That being the case, our second hypothesis (H2) is not supported.

3) *The direct effect of Aesthetics on Credibility is stronger for the female group than for the male group:* As shown in Fig. 4, the direct effects of *aesthetics* on *credibility* for the female group (0.72 | 0.70 | 0.66 | 0.67) are not significantly higher than those for the male group [0.64 | 0.65 | 0.70 | 0.49]. Neither are the total effects shown in Fig. 5b and Fig. 5c for the female group (0.80 | 0.76 | 0.72 | 0.80) significantly higher than those for the male group [0.84 | 0.80 | 0.67 | 0.73]. Therefore, our third hypothesis (H3) is not supported.

4) *The effect size of Aesthetics on Credibility is stronger for the female group than for the male group:* As clearly shown in Fig. 7, for all four webpages, the effect sizes for the female group (0.82 | 0.74 | 0.64 | 0.61) are higher than effect sizes for the male group [0.59 | 0.56 | 0.40 | 0.14]. Consequently, we conclude that our fourth hypothesis (H4) is fully supported and replicated across all UI designs.

5) *The direct effect of Usability on Credibility is higher for the male group than for the female group:* We presented the direct effects [(U→C)_{DIR}] for both genders in Fig. 4. As shown, for all four webpages, except for mobile WC, the

direct effects for the male group [0.27 | 0.20 | -0.04 | 0.27] are stronger than those for the female group (0.12 | 0.12 | 0.11 | 0.18). However, it is only the difference between the first pair of path coefficients (0.27, $p < 0.001$ and 0.12, $p < 0.05$) for mobile WA that is significant at $p < 0.05$. As a result of the non-replication of the finding across one or more UI designs, we conclude that our fifth hypothesis (H5), overall, is not fully supported. However, this calls for further investigation, especially given that three of the direct effects of *usability* on *credibility* are higher for males than for females (see Fig. 4).

6) *The effect size of Usability on Credibility is higher for the male group than for the female group*: Fig. 7 shows a plot of the effect sizes of *usability* on *credibility*. As shown, the effect sizes for the male group for all four webpages are [0.13 | 0.05 | 0.00 | 0.04], while the effect sizes for the female group are (0.02 | 0.05 | 0.00 | 0.05). While the other corresponding pairs of effect sizes are virtually equal, the first effect size (for mobile WA) for the male group (0.13—an almost medium effect size) is higher than that for the female group (0.02—a weak effect size). Based on this, we conclude that our sixth hypothesis (H6) is supported but not replicated. As such, overall, the sixth hypothesis (H6) is not fully supported.

7) *The mediation of the direct effect of Aesthetics on Credibility by Usability is higher for the grid-layout mobile design than the list-layout mobile design*: Fig. 6 shows a plot of the *usability* mediation of the effect of *aesthetics* on *credibility*. In the global, male and female models, the VAFs (a measure of mediation) for WA | WB | WC | WD are {0.14 | 0.11 | 0.00 | 0.23}, [0.24 | 0.18 | 0.00 | 0.32] and (0.10 | 0.09 | 0.00 | 0.16), respectively. As shown in the plot, at both the global and subgroup levels, the mediation by *usability* is higher for mobile WD than for mobile WA, WB and WC. In the global and male models, the *usability* mediations (0.23 and 0.32, respectively) for mobile WD are partial. However, in the female model, though there is no mediation for mobile WD, as the VAF value (0.16) is less than 0.2, the VAF value for mobile WD (0.16) is higher than those for the other webpages (0.10, 0.09 and 0.00). Consequently, given that the *usability* mediation of the effect of *aesthetics* on *credibility* is higher for mobile WD (a grid-layout design) than for the other webpages (list-layout designs), we conclude that our seventh hypothesis (H7) is supported and replicated across both genders. However, given that for the female model, the VAF value is less than 0.2, there is a need for further investigation of H7.

8) *The mediation of the direct effect of Aesthetics on Credibility by Usability is stronger for the male group than for the female group*. As shown in Fig. 6, the VAF values for mobile WA | WB | WC | WD are higher for the male group [0.24 | 0.18 | 0.00 | 0.32] than for the female group (0.10 | 0.09 | 0.00 | 0.16), except for mobile WC where there is no mediation. Therefore, our eighth hypothesis (H8) is supported and replicated across (three) UI designs.

VI. DISCUSSION

We have presented the results of our study on the interplay of *aesthetics*, *usability* and *credibility* in the judgment of mobile web designs. For quick and easy reference, we have summarized in Table IV all of our main findings, including an additional finding (H9^a), which contradicts our second hypothesis (H2). First, four of our hypotheses (H1, H4, H7 and H8) are supported and replicated across all four UI designs. Second, two of our hypotheses (H5 and H6) are supported with respect to one UI design only, but are not replicated across the other three UI designs. Third, two of the hypotheses (H2 and H3) are neither supported nor replicated. Finally, the opposite of our second hypothesis (i.e., H9^a) is supported and replicated across all four UI designs.

A. H1: Validation of Hypotheses

In this subsection, we discuss the validation and non-validation of the hypotheses summarized in Table IV.

TABLE IV. SUMMARY OF VALIDATED HYPOTHESES

No.	Hypothesis	Remark
H1	The direct effect of Aesthetics on Credibility is stronger than the direct effect of Usability on Credibility.	Supported and replicated
H2	The direct effect of Aesthetics on Usability is stronger for the female group than for the male group.	Not Supported
H3	The direct effect of Aesthetics on Credibility is stronger for the female group than for the male group.	Not Supported
H4	The effect size of Aesthetics on Credibility is higher for the female group than for the male group.	Supported and replicated
H5	The direct effect of Usability on Credibility is stronger for the male group than for the female group.	Supported but not replicated
H6	The effect size of Usability on Credibility is higher for the male group than for the female group.	Supported but not replicated
H7	The mediation of the direct effect of Aesthetics on Credibility by Usability is higher for the grid-layout mobile design than the list-layout mobile designs.	Supported and replicated
H8	The mediation of the direct effect of Aesthetics on Credibility by Usability is stronger for the male group than for the female group.	Supported and replicated
H9 ^a	The direct effect of Aesthetics on Usability is stronger for the male group than for the female group.	Supported and Replicated

Note: H9^a (opposite of H2) was not pre-formulated but turned out to be true.

1) *H1: The direct effect of Aesthetics on Credibility is stronger than the direct effect of Usability on Credibility*: As presented in Table IV, we showed through path modeling that our first hypothesis (H1) is supported by replicating it across gender and all four UI designs. This finding replicates our prior finding in our study which is based on two different cultures [7]. Thus, with the validation of this hypothesis in this

paper, we have provided more empirical evidence in the mobile web domain (which cuts across gender) by confirming that *perceived aesthetics* is a stronger determinant than *perceived usability* in predicting the *perceived credibility* of a mobile website. Moreover, this finding replicates prior findings in the web domain. For example, Fogg et al. [9] found that *design look* was the most important factor in evaluating *web credibility*. In their qualitative study among 2,684 participants, they found that *design look* was mentioned in 46.1% of participants' comments. However, it is noteworthy that while we have found *aesthetics* to be a stronger predictor of mobile *web credibility*, as compared to *usability*, our finding does not suggest that *usability* is not important (and should not be considered) when designing mobile websites. Off course, *usability* is important in the user experience, especially in the mobile domain where the relatively small screen size can pose a serious threat to the effective and efficient use of mobile applications [17]. As we have shown in the global and subgroup models (in Fig. 3 and Fig. 4), *usability* plays a role in the judgment of *credibility* as well, especially if a user has to actually use a website for some time to accomplish a task. In three of the presented mobile designs (WA, WB and WD), we have also presented empirical evidence, which shows that *usability* (though not as strongly as *aesthetics*) influences the *credibility* of a mobile website. According to Fogg et al. [83], *usability* (i.e., making sites easy to use) will equally enhance the *credibility* of a website. In summary, our finding suggests that, while ensuring that mobile web applications are designed to be easily usable given our awareness of the relatively small size of the mobile screen and the usability challenges it poses in human-computer interaction, designers should also endeavor to cater to users' hedonic needs by designing their websites to be visually appealing. This has the potential of attracting more users to their websites at the visceral level of judgment, which is based on first impression [1], [84].

2) *H2: The direct effect of Aesthetics on Usability is stronger for the female group than for the male group:* Our second hypothesis (H2), as presented in Table IV, is not supported. The influence of *aesthetics* on *usability* is aptly regarded as the "aesthetic-usability effect" [31], a cognitive bias in human-computer interaction, is also regarded to as the "halo effect" [57]. Contrary to our hypothesis, in our study the halo effect is stronger for the male group than the female group, as stated in H9^a. This is reflected in the coefficient of determination of both subgroup models (see Fig. 4), with *aesthetics* in the male model ($R^2 = [0.58 | 0.53 | 0.34 | 0.74]$) explaining the variance in *usability* more than in the female model ($R^2 = (0.40 | 0.30 | 0.32 | 0.52)$). This suggests that: (1) females' *usability perception* is more difficult to predict using *aesthetics perception* compared to males'; and (2) there are other factors (which are not captured in our model) that may account for the variance in *perceived usability*, especially for the female group. This implies that there is a need for more

research to identify and understand these other factors than *aesthetics* (e.g., *information design*, *advertising*, etc. [9]), which may impact *usability* in the mobile domain. Furthermore, the unexpected finding that *perceived aesthetics* influences *perceived usability* more strongly for the male group than for the female group suggests that the aesthetic-usability halo effect is stronger in males than in females.

3) *H3: The direct effect of Aesthetics on Credibility is stronger for the female group than for the male group:* Based on the finding that females are more likely to be motivated by intrinsic factors in their judgment than males [36], we had hypothesized that the influence of *aesthetics* on *credibility* will be stronger for the female group than for the male group (H3). However, our multigroup analysis showed no support for this hypothesis. However, there might be a need for further future research to find out whether other factors than gender (such as culture, age, etc.) may moderate the relationship between *perceived aesthetics* and *perceived credibility*.

4) *H4: The effect size of Aesthetics on Credibility is stronger for the female group than for the male group:* The result of our effect-size analysis (see Fig. 7) reveals that the effect size of *aesthetics* on *credibility* is stronger for the female group than for the male group (H4). It is interesting to note that while our third hypothesis (H3), which concerns the direct effect of *aesthetics* on *credibility*, is not supported, our fourth hypothesis (H4), which concerns the effect size of *aesthetics* on *credibility*, is supported for all four UI designs. This suggests that, in an actual setting, *perceived aesthetics* may have a stronger influence on the *perceived credibility* of websites for the female group than it does for the male group. This finding replicates the notion that females are more concerned about intrinsic factors than males in judging products, such as websites [36].

5) *H5: The direct effect of Usability on Credibility is higher for the male group than for the female group:* The fifth hypothesis (H5) is only supported with respect to one of the UI designs. Specifically, we found a significant difference between males and females with respect to mobile WA only, with *usability* having a higher effect on *credibility* for males than for females (see Fig. 4 and Table III). However, given that this finding is not replicated among any of the other web designs than mobile WA, we conclude that, overall, our hypothesis (H5) is not supported. However, given that there is evidence that supports the fifth hypothesis (H5), with respect to mobile WA, there might be a need for further investigation to be carried out to verify this hypothesis, especially given that all of the three significant *usability-credibility* relationships are higher for the male group than for the female group.

6) *H6: The effect size of Usability on Credibility is higher for the male group than for the female group:* The sixth hypothesis on the effect size of *usability* on *credibility* is only

supported with respect to mobile WA; it is not replicated across the other three UI designs. Specifically, we found that the effect size with respect to mobile WA is larger for the male group (0.12—almost medium ES) than for the female group (0.02—small ES). This is also evident in the significantly different direct effects when the male group ($\beta = 0.27$, $p < 0.001$) is compared to the female group ($\beta = 0.12$, $p < 0.05$), as indicated by bold fonts in Fig. 4. In practical terms, this suggests that, in the judgment of *credibility* with respect to mobile design, in general (or, in particular, low-aesthetic mobile web design), *usability* may matter more to males than it does to females. However, the non-replication of this finding (significant difference between males and females with respect to the $(U \rightarrow C)_{\text{DIR}}$ relationship) across two or more UI designs makes it difficult for us to conclude that our sixth hypothesis (H6) is supported. As a result, we conclude that, overall, our sixth hypothesis (H6) is not supported. However, given the evidence in favor of this hypothesis with respect to mobile WA, further investigations may need to be carried out in future research to validate or invalidate it.

7) *H7: The mediation of the direct effect of Aesthetics on Credibility by Usability is higher for the grid-layout mobile design than the list-layout mobile design:* Our seventh hypothesis states that, with respect to organization of information or content in mobile web design, *usability* will mediate the effect of *aesthetics* on *credibility* more for a grid layout (mobile WD) than for the list layouts (mobile WA, WC and WC). The reasoning behind this hypothesis is that, in a prior qualitative study, Oyibo, Ali and Vassileva [60] found that the grid-layout design is perceived as more usable than the list-layout design by both genders due to how well-spaced out the icon-based items were on the screen compared to the list-based items, which are relatively closer and almost touching. Specifically, regarding mobile WD, a participant in their study commented: “The buttons would probably be easy to press with my stubby fingers without accidentally hitting the wrong one” (p. 82) [60]. As expected (see Fig. 6), we found that, for both genders, the mediation of *usability* of the effect of *aesthetics* on *credibility* is higher for mobile WD than for the other webpages. Thus, our seventh hypothesis (H7) is supported and replicated across gender. This finding suggests that the more easy to use (e.g., navigate) a mobile website is perceived, the higher the possibility of *usability* mediating the effect of *perceived aesthetics* on *perceived credibility*.

8) *H8: The mediation of the direct effect of Aesthetics on Credibility by Usability is stronger for the male group than for the female group.* As shown in Fig. 6, the *usability* mediation of the effect of *aesthetics* on *credibility* is higher for males than for females with respect to all the webpages, except for mobile WC. Specifically, for mobile WA and WD, there is partial *usability* mediation for the male group but none for the female group. Consequently, our eighth hypothesis (H8) is supported and replicated. This suggests that males may be

more likely concerned about *usability* (an instrumental attribute) than females in the judgment of *credibility*, as found in prior research [36], [60]. However, this may need further investigation, for example, given that the hypothesis is not verified with respect to all four UIs. Aside, readers might be interested in knowing the likely reason for the lack of *usability* mediation of the effect of *aesthetics* on *credibility* with respect to mobile WC as compared to the other mobile webpages (WA, WB and WD). As a result, we provide a plausible explanation. In the global model (Fig. 5a), we see that in the $(A \rightarrow C)_{\text{IND}}$ effects for mobile {WA | WB | WC | WD}, i.e., {0.12 | 0.09 | 0.01 | 0.18}, respectively, the indirect effect for mobile WC is virtually zero (non-significant). Similarly, in the male model (Fig. 5b), we see that in the $(A \rightarrow C)_{\text{IND}}$ effects [0.21 | 0.15 | -0.03 | 0.23] for all four UIs, the indirect effect for mobile WC is virtually zero (non-significant). Finally, in the female model (Fig. 5c), we see that in the $(A \rightarrow C)_{\text{IND}}$ (0.08 | 0.07 | 0.06 | 0.13) for all four UIs, the indirect effect for mobile WC is least and non-significant as well. One possible explanation for this finding (the non-significance of $(A \rightarrow C)_{\text{IND}}$ for mobile WC) is that mobile WC was presented to all of the participants first in the online survey. It appears that, upon their initial encounter with this version of the hypothetical website (WC), the only criterion participants used in judging *credibility* was *aesthetics*; *usability*, as our path models shows was not used (i.e., *perceived usability* neither affected *perceived credibility* nor mediated the direct effect $[(A \rightarrow C)_{\text{DIR}]$, with respect to WC, as evident in the non-significant direct effect $[(U \rightarrow C)_{\text{DIR}]$ shown in Fig. 3 and Fig. 4). In other words, in the context of the elaboration likelihood model [85], both males and females used, completely, the peripheral route, i.e., the short-circuited path $[(A \rightarrow C)_{\text{DIR}]$, to judge *credibility*. This is consistent with the prior finding that the judgment of *credibility* of websites is made within the first few seconds of users’ coming in contact with the user interface and it is mainly informed by *visual appeal*. Lindgaard et al. [84] found that when a person opens a website, the first impression, predicated on the *visual appeal* of the site, is made within the first 50 milliseconds. According to Robins and Holmes [1], this first impression determines whether a person will continue using (e.g., browsing) a website or abandon it for another website. Consequently, in our study, it is very likely that the first impression, based on *visual appeal*, must have shifted participants’ attention away from usability-related attributes and concerns. Thus, with respect to mobile WC, presented first in the online survey, we find that *perceived usability* neither influences nor mediates the influence of *perceived aesthetics* on *perceived credibility*—a finding which cuts across gender.

B. Implications and Contributions

In summary, we would like to briefly discuss the overarching implication of our findings in the context of our research questions and summarize our contributions. Overall, the results of our PLSPM analysis show that, irrespective of the level of aesthetic treatment of mobile websites, the layout of

the UI design and gender, *aesthetics* is a stronger determinant of *credibility* than *usability* is. This implies that, in the mobile web domain, the hedonic need of users has to be catered to as well by designers [86], as designing websites that are only usable, but lack visual appeal, may not be able to attract a critical mass of users as a result of poor first impressions made by the websites on visitors. Therefore, mobile web designers should ensure their sites are visually appealing, as *perceived aesthetics*, at the visceral level of website evaluation, is stronger than *perceived usability* in predicting *web credibility*, as found in prior research [1], [7], [9], [87] on web design.

In conclusion, our main contributions to the body of knowledge can be summarized as follows:

1. We showed that, in the evaluation of mobile websites, *perceive aesthetics* and *perceived usability* combined can account for up to and over 50% of the variance in *perceived credibility*.
2. We showed that, regardless of gender and layout of UI design, *perceived aesthetics* is stronger than *perceived usability* in predicting the *perceived credibility* of a mobile website.
3. We presented new empirical evidence regarding the moderating effect of gender and design layout in the interrelationships among the three web design constructs of interest. They include the following:
 - a. The *aesthetic-usability* effect is stronger for males than for females.
 - b. The effect size of *aesthetics* on *credibility* is larger for females than for males.
 - c. The *usability* mediation of the effect of *aesthetics* on *credibility* is larger for males than for females.
 - d. The *usability* mediation of the effect of *aesthetics* on *credibility* is larger for a grid-layout than for a list-layout mobile website.

Finally, our study, to the best of our knowledge, is the first of its kind in the mobile web domain that has presented empirical evidence on the interrelationships among the three important web design constructs (*aesthetics*, *usability* and *credibility*) in a replicated fashion across different UI designs and gender, using a multicultural population sample that cuts across five continents. This opens up the possibility of a broader generalization of findings.

VII. LIMITATION AND FUTURE WORK

The main limitation of our study is that the measurement of the three web design constructs (*perceived aesthetics*, *perceived usability* and *perceived credibility*), which we investigated, was based on perception and not actual use of the mobile websites. This might have impacted our findings, thereby limiting the scope of our findings to the visceral level of perception, as our current findings may differ in the context of actual use of the mobile websites. A second limitation of our study is that, for convenience in answering the online survey questions, most participants might have viewed the mobile web

designs on a desktop rather than a mobile device. This might have affected their perception of the mobile web designs and the way they responded. Finally, the fixed order of presentation of the mobile web designs in the online survey must have also impacted our findings to a certain degree, as we pointed out when discussing the eighth hypothesis. However, in future work, we look forward to addressing some of these limitations, including investigating the qualitative response of users to the four web designs covered in this paper and how it may differ across culture, age, gender and UI design.

VIII. CONCLUSION

In this paper, we presented a partial least square path model of the interrelationships among three key web design constructs of mobile websites. Specifically, we attempt to uncover which of the key determinants of *web credibility* (*perceived aesthetics* or *perceived usability*) has a stronger influence on *perceived credibility* in the mobile domain, and the role UI design and gender play. Using a multicultural sample of 526 participants from five continents (Africa, Asia, North America, South America and Europe) and four systematically designed mobile websites, our path model reveals that, irrespective of UI design and gender, *perceived aesthetics* has a stronger effect than *perceived usability* on the *perceived credibility* of a mobile website. Secondly, our model shows that the effect size of *perceived aesthetics* on *perceived credibility* is larger for females than for males. Thirdly, our model shows that the *aesthetic-usability* effect and the *usability* mediation of the effect of *perceived aesthetics* on *perceived credibility* are stronger for males than for females. Finally, our model shows that the *usability* mediation of the direct effect between *aesthetics* and *credibility* may depend on the layout of content on the screen of the mobile device. Specifically, there is a stronger *usability* mediation of the effect of *aesthetics* on *credibility* for the grid-layout mobile design than the list-layout mobile designs. This suggests that users may be more cognizant of *usability* concerns in the judgment of *credibility* in a grid-layout mobile design than in a list-layout mobile design. However, this finding requires further investigation.

The overall implication of our findings is that, while *usability* (a utilitarian attribute) is of great concern in the mobile domain given the relatively small size of the mobile device, designers should not lose sight of the need to satisfy the hedonic needs of users as well, as users, as we have shown, perceive *aesthetics* (*visual appeal*) as very important in the whole user experience and in the overall judgment of *web credibility*, especially at the visceral (non-cognitive) level of perception. In conclusion, our findings suggest that, although, in some specific use cases, one of the two web design attributes may be given priority over the other, in general, there is a need for designers to strike a balance between hedonic and utilitarian concerns [88], while specifically paying close attention to *visual aesthetics*—which the user first comes in contact with in human-computer interaction—and gender differences as well.

ACKNOWLEDGEMENT

The authors would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC) for providing the funding which made this research possible.

REFERENCES

- [1] D. Robins and J. Holmes, "Aesthetics and credibility in web site design," *Information Processing & Management*, vol. 44, no. 1, pp. 386–399, 2008.
- [2] B. Fogg, "What Makes a Website Credible?," *Stanford University*, 2003. [Online]. Available: http://static.lukew.com/web_credibility_lecture.pdf. [Accessed: 17-Feb-2017].
- [3] Trust Verified, "The Importance of Website Credibility," *Trust Verified*. [Online]. Available: <http://trust-verified.org/the-importance-of-website-credibility/>. [Accessed: 17-Feb-2017].
- [4] B. J. Fogg, J. Marshall, O. Laraki, A. Osipovich, C. Varma, N. Fang, J. Paul, A. Rangnekar, J. Shon, P. Swani, M. Treinen, and C. Hall, "What Makes Web Sites Credible? A Report on a Large Quantitative Study CHI 2001," in *Sigchi '01*, 2001, pp. 61–68.
- [5] T. Lavie and N. Tractinsky, "Assessing dimensions of perceived visual aesthetics of web sites," *International Journal of Human-Computer Studies*, vol. 60, no. 3, pp. 269–298, Mar. 2004.
- [6] C.-H. Liu, C.-F. Lee, and H.-L. Lee, "The Influence of HCI Design Aesthetics on Website Creditability - Using on Online Banking Website Interfaces as an Example," in *Design Computing and Cognition*, 2010, pp. 1–20.
- [7] K. Oyibo and J. Vassileva, "The Interplay of Aesthetics, Usability and Credibility in Mobile Websites and the Moderation by Culture," in *Brazilian Symposium on Human Factors in Computer Systems*, 2016.
- [8] A. Sonderegger, A. Uebelbacher, M. Pugliese, and J. Sauer, "The Influence of Aesthetics in Usability Testing: The Case of Dual-domain Products," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 21–30, 2014.
- [9] B. J. Fogg, C. Soohoo, D. R. Danielson, L. Marable, J. Stanford, and E. R. Tauber, "How do users evaluate the credibility of Web sites?: a study with over 2,500 participants," in *Proceedings of the 2003 conference on Designing for user experiences*, 2003, pp. 1–15.
- [10] D. Cyr, "Gender and website design across cultures," *17th European Conference on Information Systems*, pp. 279–291, 2009.
- [11] N. Tractinsky, "Aesthetics and apparent usability: Empirically assessing cultural and methodological issues," *Conference on Human Factors in Computing Systems*, pp. 115–122, 1997.
- [12] Z. Göcza, "Myth # 25 : Aesthetics are not important if you have good usability," [Online]. Available: <http://uxmyths.com/post/1161244116/myth-25-aesthetics-are-not-important-if-you-have-good-us>. [Accessed: 02-Apr-2016].
- [13] N. Tractinsky, A. S. Katz, and D. Ikar, "What is beautiful is usable," *Interacting with computers*, vol. 13, no. 2, pp. 127–145, 2000.
- [14] M. Hassenzahl and A. Monk, "The Inference of Perceived Usability From Beauty," *Human-Computer Interaction*, vol. 24, no. March 2015, pp. 37–41, 2010.
- [15] A. David and P. Glore, "The Impact of Design and Aesthetics on Usability, Credibility, and Learning in Online Courses," *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2010*, no. 2004, p. 42, 2010.
- [16] N. E. Youngblood and J. MacKiewicz, "A usability analysis of municipal government website home pages in Alabama," *Government Information Quarterly*, vol. 29, no. 4, pp. 582–588, 2012.
- [17] M. Chae and J. Kim, "Do size and structure matter to mobile users? An empirical study of the effects of screen size, information structure, and task complexity on user activities with standard web phones," *Behaviour & Information Technology*, vol. 23, no. 3, pp. 165–181, 2004.
- [18] N. Kumar, "Usability for the new PC: mobile devices," *Third Annual Symposium Theory and Research in HCI*, 2012.
- [19] Interaction Design Foundation, "The 7 Factors that Influence User Experience." [Online]. Available: https://www.interaction-design.org/literature/article/the-7-factors-that-influence-user-experience?utm_content=buffer4f582&utm_medium=social&utm_source=facebook.com&utm_campaign=buffer. [Accessed: 27-Dec-2016].
- [20] S. Mahlke and M. Thüring, "Studying antecedents of emotional experiences in interactive contexts," *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*, pp. 915–918, 2007.
- [21] T. Lavie and N. Tractinsky, "Assessing dimensions of perceived visual aesthetics of web sites," *International Journal of Human Computer Studies*, vol. 60, no. 3, pp. 269–298, 2004.
- [22] N. Tractinsky, "Aesthetics and apparent usability: empirically assessing cultural and methodological issues," in *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems*, 1997, pp. 115–122.
- [23] K. Oyibo and J. Vassileva, "What Drives Perceived Usability in Mobile Web Design : Classical or Expressive Aesthetics?," in *19th International Conference on Human-Computer Interaction*, 2017, pp. 445–462.
- [24] D. A. Norman, *The psychology of everyday things*. Basic Books, 1988.
- [25] M. Kurosu and K. Kashimura, "Apparent Usability vs. Inherent Usability: Experimental Analysis on the Determinants of the Apparent Usability," in *Conference Companion on Human Factors in Computing Systems*, 1995, pp. 292–293.
- [26] O. Taebi, H. Aldabbas, and M. Clarkson, "Designing for Usability or Aesthetics in E-commerce Websites," 2012.
- [27] J. Nielsen, *Usability Engineering*, vol. 25, no. 3. Elsevier, 1993.
- [28] M. Kurosu, K. Kashimura, and O. F. Creativity, "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, 1995.
- [29] K. Reinecke, T. Yeh, L. Miratrix, R. Mardiko, Y. Zhao, J. Liu, and K. Z. Gajos, "Predicting users' first impressions of website aesthetics with a quantification of perceived visual complexity and colorfulness," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*, pp. 2049–2058, 2013.
- [30] K. E. Kripintiris, "Web aesthetics and usability: An empirical evaluation of white space," *ProQuest Dissertations and Theses*, p. 65, 2008.
- [31] W. Lidwell, K. Holden, and J. Butler, "Universal Principles of Design," *Universal principles of design: 125 ways to enhance usability, influence perception, increase appeal, make better design decisions, and teach through design.*, pp. 1–271, 2010.
- [32] D. Cyr and C. Bonanni, "Gender and website design in e-business," *International Journal of Electronic Business*, vol. 3, no. 6, p. 565, 2005.
- [33] S. Ferebee, "The influence of gender and involvement level on the perceived credibility of web sites," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5033 LNCS, pp. 279–282, 2008.
- [34] S. J. Simon, "The impact of culture and gender on web sites: an empirical study," *SIGMIS Database*, vol. 32, no. 1, pp. 18–37, 2001.
- [35] J. Meyers-Levy, "Gender Differences in Information Processing: A Selectivity Interpretation," *Cognitive and Affective Responses to Advertising*, pp. 219–260, 1989.
- [36] Y. Sun, K. H. Lim, C. Jiang, J. Z. Peng, and X. Chen, "Do males and females think in the same way? An empirical investigation on the gender differences in Web advertising evaluation," *Computers in Human Behavior*, vol. 26, no. 6, pp. 1614–1624, 2010.
- [37] K. Oyibo, Y. S. Ali, and J. Vassileva, "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*, 2016, pp. 75–84.
- [38] A. Harley, "Trustworthiness in Web Design: 4 Credibility Factors," *Nielsen Norman Group (NN/g)*, 2016. [Online]. Available: <https://www.nngroup.com/articles/trustworthy-design>.
- [39] D. Cyr and H. Trevor-Smith, "Localization of web design: An empirical comparison of German, Japanese, and United States web site characteristics," *Journal of the American Society for Information Science and Technology*, vol. 55, no. 13, pp. 1199–1206, 2004.
- [40] B. J. Fogg, J. Marshall, O. Laraki, A. Osipovich, C. Varma, N. Fang, J. Paul, A. Rangnekar, J. Shon, P. Swani, M. Treinen, and C. Hall, "What Makes Web Sites Credible? A Report on a Large Quantitative Study CHI 2001," pp. 61–68, 2001.
- [41] D. S. Soper, "User Interface Design and the Halo Effect: Some Preliminary Evidence," in *Twentieth Americas Conference on Information Systems, Savannah*, 2014, pp. 1–11.

- [42] K. Reinecke and A. Bernstein, "Improving performance, perceived usability, and aesthetics with culturally adaptive user interfaces," *ACM Transactions on Computer-Human Interaction*, vol. 18, no. 2, pp. 1–29, 2011.
- [43] B. J. Fogg, *Persuasive Technology: Using Computers to Change What We Think and Do*. Morgan Kaufmann, 2003.
- [44] B. J. Fogg, "Prominence-interpretation theory: Explaining how people assess credibility online," *Conference on Human Factors in Computing Systems - Proceedings*, pp. 722–723, 2003.
- [45] C. Sartwell, "Beauty," in *The Stanford Encyclopedia of Philosophy*, Winter 201., E. N. Zalta, Ed. Metaphysics Research Lab, Stanford University, 2016.
- [46] E. Burns Coleman, C. Hartney, and Z. Alderton, "Defining 'Social Aesthetics,'" *Aesthetics*, vol. 23, no. 1, pp. 1–11, 2013.
- [47] H. Caygill, "Aesthetics and civil society: theories of art and society 1640 - 1790," University of Sussex, 1982.
- [48] L. Mastin, "The Basics of Philosophy," 2008. [Online]. Available: http://www.philosophybasics.com/branch_aesthetics.html. [Accessed: 23-Sep-2017].
- [49] D. Burnham, "Immanuel Kant: Aesthetics," *Internet Encyclopedia of Philosophy*. [Online]. Available: <http://www.iep.utm.edu/kantaest/>. [Accessed: 12-Nov-2017].
- [50] Kanaan, "Philosophy's Moods: The Affective Grounds of Thinking," *Igarss 2014*, no. 1, pp. 1–5, 2014.
- [51] N. Zangwill, "Aesthetic Judgment," in *The Stanford Encyclopedia of Philosophy*, Fall 2014., E. N. Zalta, Ed. Metaphysics Research Lab, Stanford University, 2014.
- [52] N. Tractinsky, "Visual Aesthetics," *Interaction-Design*, pp. 1–71, 2002.
- [53] A. Sonderegger, J. Sauer, and J. Eichenberger, "Expressive and classical aesthetics: two distinct concepts with highly similar effect patterns in user–artifact interaction," *Behaviour & Information Technology*, vol. 33, no. 11, pp. 1180–1191, 2014.
- [54] N. Tractinsky, "Toward the study of aesthetics in information technology," *ICIS 2004 Proceedings*, p. 62, 2004.
- [55] V. Venkatesh and F. D. Davis, "A theoretical extension of the technology acceptance model: Four longitudinal field studies," *Information Systems Research*, vol. 46, no. 2, pp. 186–204, 2000.
- [56] K. A. Butler, "Usability Engineering Turns 10," *Interactions*, vol. 3, no. 1, pp. 58–75, 1996.
- [57] A. Sonderegger and J. Sauer, "The influence of design aesthetics in usability testing: Effects on user performance and perceived usability," *Applied Ergonomics*, vol. 41, no. 3, pp. 403–410, 2010.
- [58] M. Hassenzahl, "The interplay of beauty, goodness, and usability in interactive products," *Human-Computer Interaction*, vol. 19, no. 4, pp. 319–349, 2004.
- [59] Y. M. Li and Y. S. Yeh, "Increasing trust in mobile commerce through design aesthetics," *Computers in Human Behavior*, vol. 26, no. 4, pp. 673–684, 2010.
- [60] K. Oyibo, Y. S. Ali, and J. Vassileva, "Gender Difference in the Credibility Perception of Mobile Websites: A Mixed Method Approach," in *User Modeling, Adaptation and Personalization (UMAP 2016)*, 2016, pp. 1–10.
- [61] Y.-M. Li and Y.-S. Yeh, "Increasing trust in mobile commerce through design aesthetics," *Computers in Human Behavior*, vol. 26, no. 4, pp. 673–684, 2010.
- [62] K. Oyibo and J. Vassileva, "The Interplay of Aesthetics, Usability and Credibility in Mobile Websites and the Moderation by Culture," 2016.
- [63] Y. Neumark, L. Flum, C. Lopez-Quintero, and R. Shtarkshall, "Quality of online health information about oral contraceptives from Hebrew-language websites," *Israel Journal of Health Policy Research*, vol. 1, no. 1, pp. 1–10, 2012.
- [64] M. Miche and M. Noirhomme-frature, "Gender differences in graphic design for the Web Categories and Subject Descriptors," *Revista Eletrônica de Sistemas de Informação*, vol. 5, no. 1, 2009.
- [65] G. Moss and R. Gunn, "Gender differences in website design: Implications for education," *Int Conf on Cybernetics and Information Technologies, Systems and Applications/Int Conf on Computing, Communications and Control Technologies, Vol 1*, pp. 54–59, 2007.
- [66] S. Djamasbi, T. Tullis, J. Hsu, E. Mazuera, K. Osberg, and J. Bosch, "Gender preferences in web design: usability testing through eye tracking," *Proceedings of the Americas Conference on Information Systems*, pp. 1–8, 2007.
- [67] K. Oyibo, Y. S. Ali, and J. Vassileva, "An empirical analysis of the perception of mobile website interfaces and the influence of culture," in *Proceedings of the Personalization in Persuasive Technology Workshop, Persuasive Technology 2016*, 2016, vol. 1582, pp. 44–56.
- [68] T. Ramakrishnan, V. Prybutok, and D. A. Peak, "The moderating effect of gender on academic website impression," *Computers in Human Behavior*, vol. 35, pp. 315–319, 2014.
- [69] L. Xue and C. C. Yen, "Towards female preferences in design - A pilot study," *International Journal of Design*, vol. 1, no. 3, pp. 11–27, 2007.
- [70] M. Arcand and J. Nantel, "What drives ease of use perceptions and behavioral intentions towards websites for men and women? A study of French-Canadian Internet Users," *Management and Avenir*, vol. 2, no. 32, pp. 154–172, 2010.
- [71] E. Shinnars, "Effects of The 'What is Beautiful is Good' Stereotype on Perceived Trustworthiness," *Shinnars UW-L Journal of Undergraduate Research XII*, pp. 1–5, 2009.
- [72] K. Oyibo, Y. S. Ali, and J. Vassileva, "An empirical analysis of the perception of mobile website interfaces and the influence of culture," in *CEUR Workshop Proceedings*, 2016, vol. 1582, pp. 44–56.
- [73] P. van Schaik and J. Ling, "The role of context in perceptions of the aesthetics of web pages over time," *International Journal of Human-Computer Studies*, vol. 67, no. 1, pp. 79–89, 2009.
- [74] A. De Angeli, A. Sutcliffe, and J. Hartmann, "Interaction, usability and aesthetics: What influences users' preferences?," in *Proceedings of the 6th Conference on Designing Interactive Systems*, 2006, no. November, pp. 271–280.
- [75] S. P. S. Setterstrom, "Assessing credibility and aesthetic perception across different exposure times on a health care information website."
- [76] L. Bergkvist and J. R. Rossiter, "The Predictive Validity of Multiple-Item Versus Single-Item Measures of the Same Constructs," *Journal of Marketing Research*, vol. 44, no. 2, pp. 175–184, 2007.
- [77] L. Bergkvist, "Appropriate use of single-item measures is here to stay," *Marketing Letters*, vol. 26, no. 3, pp. 245–255, 2015.
- [78] G. Sanchez, "PLS Path Modeling with R," *Berkley: Trowchez Editions*, 2013.
- [79] J. F. Hair, G. T. M. Hult, C. M. Ringle, and M. Sarstedt, *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. Sage Publications, Inc, Washington DC, 2014.
- [80] K. J. Preacher and K. Kelley, "Effect size measures for mediation models: quantitative strategies for communicating indirect effects," *Psychological Methods*, vol. 16, no. 2, pp. 93–115, 2011.
- [81] J. Cohen, "Statistical power analysis for the behavioral sciences," *Statistical power analysis for the behavioral sciences (rev. ed.)*, vol. 2nd, 1988.
- [82] J. A. Durlak, "How to select, calculate, and interpret effect sizes," *Journal of Pediatric Psychology*, vol. 34, no. 9, pp. 917–928, 2009.
- [83] J. A. Jacko and A. Sears, "The human-computer interaction handbook: fundamentals, evolving technologies, and emerging applications," *The human-computer interaction handbook*, p. 1277, 2003.
- [84] G. Lindgaard, G. Fernandes, C. Dudek, and J. Brown, "Attention web designers: You have 50 milliseconds to make a good first impression!," *Behaviour & Information Technology*, vol. 25, no. 2, pp. 115–126, 2006.
- [85] R. E. Petty, J. T. T. Cacioppo, and R. E. Nisbett, "The elaboration likelihood model of persuasion," *Advances in experimental social psychology*, vol. 19, no. 1, pp. 123–205, 1986.
- [86] M. Hassenzahl, "The Thing and I: Understanding the Relationship Between User and Product," *Funology*, vol. 3, pp. 31–42, 2003.
- [87] K. Oyibo, R. Orji, and J. Vassileva, "The Influence of Personality on Mobile Web Credibility," in *Adjunct Proceedings of User Modeling, Adaptation and Personalization (UMAP 2017)*, 2017.
- [88] U. K. Yusof, L. K. Khaw, H. Y. Ch'ng, and B. J. Neow, "Balancing between usability and aesthetics of Web design," *Proceedings 2010 International Symposium on Information Technology - Visual Informatics, ITSIM'10*, vol. 1, 2010.