What Drives the Perceived Credibility of Health Apps: Classical or Expressive Aesthetics?

Kiemute Oyibo
University of Saskatchewan
S7N 5C9, Saskatoon, Canada
kiemute.oyibo@usask.ca

Ifeoma Adaji
University of Saskatchewan
S7N 5C9, Saskatoon, Canada
ifeoma.adaji@usask.ca

Julita Vassileva
University of Saskatchewan
S7N 5C9, Saskatoon, Canada
jiv@cs.usask.ca

ABSTRACT

Research has shown that *visual aesthetics* is a key determinant of the *perceived credibility* of human-computer-interaction systems. However, there is limited research on which of the two dimensions of *visual aesthetics* (*classical* and *expressive*) has a stronger impact on the *perceived credibility* of persuasive health apps. Consequently, we conducted an empirical study among 669 subjects to investigate: (1) which of the two dimensions has a stronger influence on the *perceived credibility* of fitness apps modeling exercise behavior; and (2) the moderating effect of gender-based personalization. Our results show that, regardless of gender-based personalization, it is the perception of *classical aesthetics* that determines the judgment of the *credibility* of a persuasive health app. *Expressive aesthetics* has no significant influence on *perceived credibility*. Our findings underscore the need for designers of persuasive systems in the health domain to focus more on the *classical* dimension of *aesthetics* (orderliness, clarity and simplicity) when designing health applications for behavior change in order to enhance their *perceived credibility*.

CCS CONCEPTS

• **Human-centered computing** → Human computer Interaction → HCI design and evaluation methods → User models

KEYWORDS

Classical aesthetics; expressive aesthetics, credibility; fitness app; persuasive technology; gender, personalization; path model

ACM Reference format:

K. Oyibo, I. Adaji, and J. Vassileva. 2018. What Drives the Perceived Credibility of Health Apps: Classical or Expressive Aesthetics? In *Proceedings of the Third International Workshop on Health Recommender Systems co-located with Twelfth ACM Conference on Recommender Systems (HealthRecSys'18), Vancouver, BC, Canada, October 6, 2018*, 6 pages.

HealthRecSys'18, October 6, 2018, Vancouver, BC, Canada.

1 INTRODUCTION

Credibility is an important design attribute that determines the adoption and usage of persuasive systems that are designed to promote behavior change [1], [2]. Generally, in human-computer interaction (HCI) systems design, research [3]–[5] has shown that *visual aesthetics* plays a very important role in the judgment of *credibility*. While there has been a substantial amount of research on the influence of *visual aesthetics* as a higher-order construct on *(perceived) credibility* (e.g., [3]–[7]), in the health domain, research investigating which of its two key dimensions (*classical* and *expressive*) has a stronger impact on *credibility* is scarce. Uncovering the more important dimension in the relationship between *visual aesthetics* and *perceived credibility* will help designers know which to focus more on when designing persuasive systems in the health domain, which users can trust and use. As Robins and Holmes observed, "*If a web site is not perceived as credible, it is unlikely it will be used*" (p386) [7].

Consequently, we carried out an empirical study in the fitness domain among 669 participants, using exercise behavior modeling as a case study. Behavior modeling is a persuasive strategy used in persuasive systems (such as fitness apps) to motivate behavior change. In behavior modeling, "*an expert shows [a] person how to correctly perform a behavior, for example, in class or on video*" (p. 382) [8]. The expert, also known as a social agent or coach, can be real or virtual. However, in computerized systems, for users to imitate the modeled behavior demonstrated by the virtual coach, there is need for the users to have trust in its visual design. This has become important due to the fact that a badly modeled exercise behavior can lead to a body injury when the user tries to replicate it. Thus, the behavior models have to be designed in a such a way that enhances their *perceived credibility*.

Based on our structural equation model (SEM), our results show that it is solely the perception of the *classical* dimension of *aesthetics* that determines the *perceived credibility* of the health app. *Expressive aesthetics* has no significant influence on *perceived credibility*. Our finding underscores the need for designers of persuasive health systems, such as behavior models in fitness apps, to focus more on *classical aesthetics* (cleanness, orderliness, clarity and pleasantness) than *expressive aesthetics* as a means to increase the *perceived credibility* of their persuasive applications.

The rest of this paper is organized as follows. Section 2 focuses on background and related work. Section 3 focuses on the research method. Section 4 focuses on the result. Section 5 focuses on the discussion. Finally, Section 6 focuses on the conclusion.

2 BACKGROUND AND RELATED WORK

This section provides a brief overview of the concepts of *aesthetics* and *credibility*, their relationships, and a review of related work.

2.1 Aesthetics

Aesthetics refers to sensory pleasure and delight derived from the perception of an object or artifact. It can be either a property of an object or perceived, or both [9]. Specifically, in this paper, we are concerned with the subjective judgment of *aesthetics*, which is in the eye of the beholder (i.e., the user). In HCI, *visual aesthetics*—defined as the pleasing appearance of HCI artifacts [10], user interfaces and systems—is regarded as a multi-dimensional construct composed of *classical aesthetics* and *expressive aesthetics* [11]. *Classical aesthetics* entails the historic notion of beauty: "*simplicity*," "*orderliness*," *proportion*, "*symmetry*," etc. These attributes of *classical aesthetics* highly correlate with *perceived usability* [11], [12]. Thus, it is described by words such as "*well-organized*," "*clear*," "*clean*," etc. On the other hand, *expressive aesthetics* has to do with the expressive power of the HCI system designer, which reflects his/her creativity and originality. Thus, it is described by words such as "*original*," "*creative*," "*fascinating*," "*sophisticated*," etc. [11].

2.2 Credibility

Credibility refers to the believability of a HCI system [1]. Just like the subjective notion of *aesthetics*, *credibility* is perceived by the user and thus is not a property of the system. According to Fogg [13], the *perceived credibility* of a HCI system is based on the simultaneous judgment of the *perceived trustworthiness* of the system and the *perceived expertise* of the designer of the system.

2.3 Aesthetics and Credibility Relationship

A substantial amount of research has been carried out with respect to the relationship between *perceived aesthetics* and *perceived credibility* in the evaluation of HCI systems. Robins and Holmes [7] examined the link between the *perceived aesthetics* and the *perceived credibility* of websites. They found that a *high-aesthetic* website is more likely to be perceived *credible* than a *low-aesthetic* website. Alsudani and Casey [14] conducted a study on the judgment of the *credibility* of recruitment agency websites. They found that *aesthetics* characteristics such as *unity in design*, *balance* and *harmony* significantly influence *perceived credibility*. Fogg et al. [15] conducted a large scale study on the judgment of *web credibility*. They found that about 46% of participants' comments on the *credibility* of various websites cutting across different domains were about *design look*—which is related to *visual aesthetics* [4].

Oyibo and Vassileva [4], [5] investigated which of *perceived aesthetics* and *perceived usability* has a stronger effect on *perceived credibility* in the tourism domain. They found that *perceived aesthetics* has a stronger effect than *perceived usability* on *perceived credibility*. Furthermore, Oyibo et al. [6] investigated which of the two dimensions of *perceived aesthetics* has a stronger effect on *perceived credibility* in the same domain. They found that, irrespective of age and gender, the *classical* dimension has a stronger effect than the *expressive* dimension on *perceived credibility*.

However, they are yet to be studies in the health domain focused on uncovering the more influential dimension of *visual aesthetics* on the *perceived credibility* of persuasive systems. The reviewed studies in the first paragraph did not address this research question. Moreover, Oyibo et al.'s [6] findings in the second paragraph are in the tourism domain, which may not generalize to the health domain. Thus, the main thrust of this paper is to investigate which of the two dimensions of *visual aesthetics* has a stronger influence on *perceived credibility*. The secondary objective of this paper is to investigate whether the previous finding in the tourism domain (*classical aesthetics has a stronger influence than expressive aesthetics on perceived credibility*) generalizes to the health domain.

3 METHOD

This section focuses on our research question, research model, measurement scales and participants' demographics.

3.1 Research Objective

The aim of this study is to answer the following research questions on the relationship between the two *aesthetics* dimensions and *perceived credibility* in the health domain:

1. *What drives the perceived credibility of fitness apps featuring exercise behavior models: classical aesthetics or expressive aesthetics, or both?*
2. *Are the interrelationships among all three design constructs moderated by the gender of the behavior model and the gender of the observer?*
3. *Does the finding that "classical aesthetics has a stronger influence than expressive aesthetics on perceived credibility in the tourism domain"* [6] *generalize to the health domain?*

3.2 Research Design

To address our research questions, we designed a fitness app prototype featuring exercise behavior models. We considered race (black/white), gender (male/female) and exercise-type (push-up/squat) in the design of the behavior models (Figure 1).[1] This resulted in eight versions. Push-up and squat were chosen because they are among the most commonly featured exercise-types in the fitness apps market. We randomly presented one of the eight designs to each participant and asked them to answer questions on *classical* and *expressive aesthetics*, and *perceived credibility*.

[1] In this paper, we are specifically investigating the moderating effect of the gender of the behavior model as well as the gender of the observer.

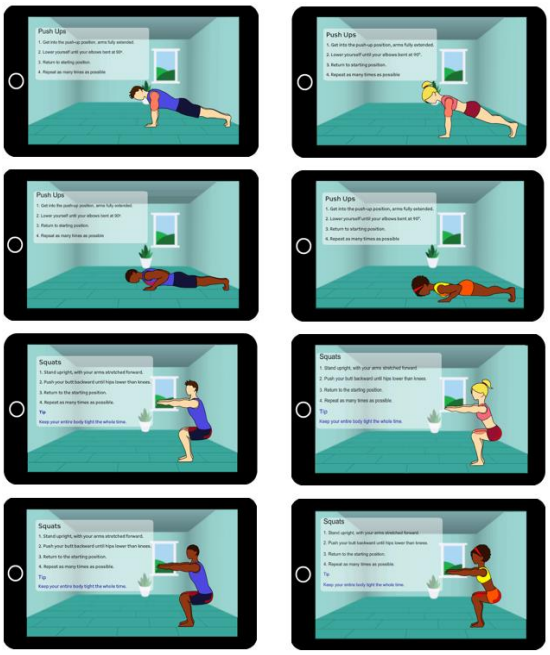

Figure 1. Behavior models performing push-up/squat [16]

3.3 Measurement Instruments

To measure the three constructs of interest, we used existing validated instruments. For *classical* and *expressive aesthetics*, we used the scale proposed by Lavie and Tractinsky [11], as adapted by van Schaik and Ling [17]. For *credibility*, we used a single-scale item scale, which research has shown is as reliable as multi-item scales [18]. All three scales ranged from "*Strongly Disagree (1)*" to "*Strongly Agree (7).*" Table 1 shows their respective items. In the questionnaire, the overarching question preceding all of the items (randomized) is, "*Please rate the visual design above on the following criteria based on your first impression.*"

Table 1: Visual design construct's scales and items

Measure	Items in each Scale
Classical Aesthetics	(1) The design is visual. (2) The design is clean. (3) The design is pleasant.
Expressive Aesthetics	(1) The design is creative. (2) The design is sophisticated. (3) The design is fascinating.
Perceived Credibility	The design is credible
Participant's Comment	Please kindly tell us the impression the visual design above had on you.

3.4 Research Model and Hypotheses

Based on prior literature, we formulated three hypotheses, which are represented in the research model shown in Figure 2. All three hypotheses (H1, H2 and H3) are based on the findings by Oyibo et al. [6] in the tourism domain. Specifically, they found that *classical aesthetics* has a stronger influence than *expressive aesthetics* on *perceived credibility*. However, in the health domain, we adopted an exploratory approach to investigate which of the two *aesthetics* dimensions has a stronger influence on *perceived credibility*. We also used an exploratory approach to investigate whether the three interrelationships in Figure 2 are moderated by the gender of the behavior model and the gender of the observer.

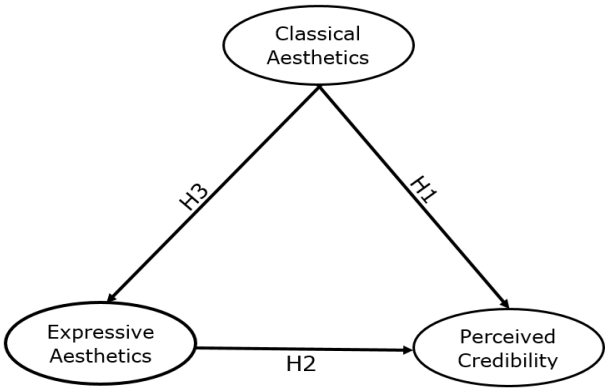

Figure 2. Research model

3.5 Participants

The survey was approved by the Behavioral Research Ethics Board of the University of Saskatchewan. Thereafter, it was posted on Amazon Mechanical Turk to recruit participants resident in Canada and United States. Each of the participants who completed the survey was compensated with $0.6 for their time. Table 2 shows the demographics of the valid participants after cleaning: 48.9% females and 51.1% males.

Table 2: demographics of participants (n = 669)

Criterion	(Female, Male) = (327, 342)
Age	18-24 (56, 70); 25-34 (139, 157); 35-34 (79, 76); 45-54 (38, 22); 54+ (15, 17)
Education	Technical/Trade School (47, 39); High School (66, 70); BSc (154, 162); MSc (42, 54); PhD (9, 6); Others (9, 11)
Country of Origin	Canada (104, 111); United States (194, 183); Others (26, 51)
Years on the Internet	0-3 (2, 2); 4-6 (18, 13); 7-9 (20, 40); 10+ (287, 287)
Behavior Model Distribution	Male behavior models (177, 175); Female behavior models (150, 167)

4 RESULT

In this section, we present our measurement models, structural models, and multigroup analysis results obtained from the SEM analysis using "PLSPM" package in R [19].

4.1 Measurement Models

We evaluated the inner models to ensure the reliability of our structural models [20]. Specifically, the indicator reliability metric

(outer loadings) and internal consistency reliability metric (DG.rho) were greater than 0.7. Similarly, the convergent validity (Average Variance Extracted) was greater than 0.5. Finally, with respect to discriminant validity, no construct's indicator loaded higher on any other construct than itself.

4.2 Structural Models

Figure 3 shows the global models for the male and female observers. The goodness of fit (GOF) represents how well the model fits the data, which is moderate: 56% for males and 66% for females. The coefficient of determination (R^2) represents the amount of variance of *perceived credibility* explained by the model: 43% for males and 56% for females—also moderate. The path coefficient (β) represents the direct effect of one construct on another. Overall, the direct effect of *classical aesthetics* on *perceived credibility* is significant for males (β = 0.58, p < 0.001) and for females (β = 0.65, p < 0.001). However, the direct effect of *expressive aesthetics* on *perceived credibility* is not significant for both males (β = 0.12, p = n.s) and females (β = 0.13, p = n.s). Moreover, we built subgroup models based on the following criteria (dyads):

1. *Males* evaluating *male* behavior models (MM)
2. *Males* evaluating *female* behavior models (MF)
3. *Females* evaluating *female* behavior models (FF)
4. *Females* evaluating *male* behavior models (FM)

MM and FF represent gender-based personalization, while MF and FM represent contra-tailoring. As shown in Figure 4, the βs, GOFs and R^2s, to a large extent, confirm the findings at the global level. For example, in the global model, we see that the relationship between *classical aesthetics* and *perceived credibility* is significant (p < 0.001), while that between *expressive aesthetics* and *perceived credibility* is not significant. This is perfectly confirmed at the subgroup level, with the relationship between *classical aesthetics* and *perceived credibility* for FM being the strongest (β = 0.82, p < 0.001) and that for MM being the weakest (β = 0.46, p < 0.001).

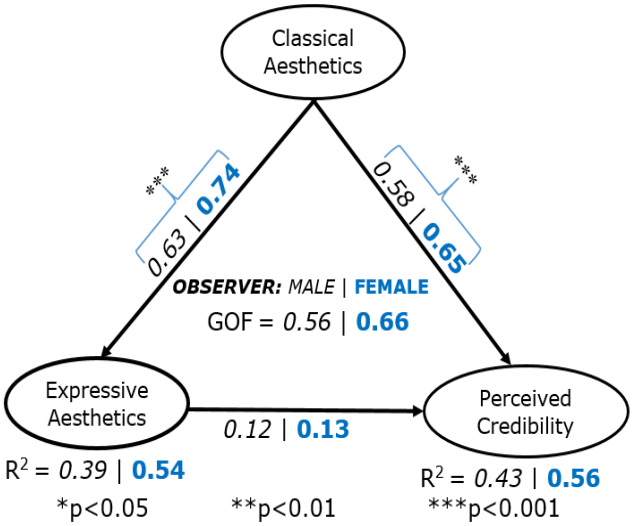

Figure 3. Global path model for male and female observers

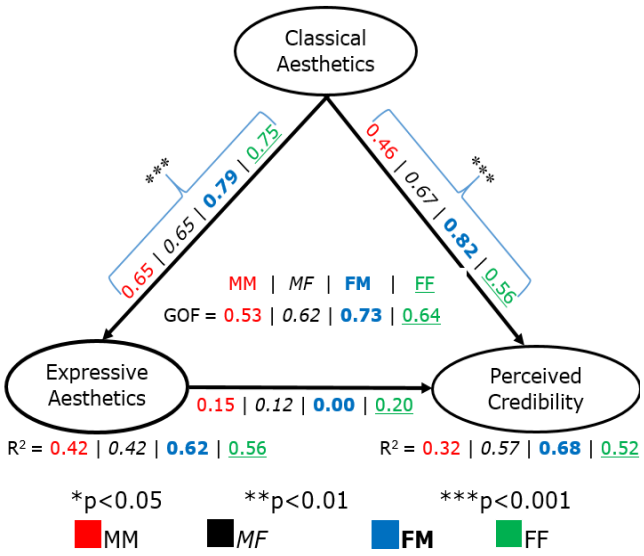

Figure 4. Path models based on gender of observer (left) and gender of behavior model (right) – M = Male, F = Female

4.3 Multigroup Analysis

To uncover how the gender of the observer and that of the behavior model moderates the interrelationships among the three design constructs, we conducted a multigroup analysis (MGA) as shown in Table 3.

Table 3: Multigroup analysis based on gender

Relationship	**M**	**F**	***p***
Classical Aesthetics→Credibility	.58	.65	n.s
Expressive Aesthetics→Credibility	.12	.13	n.s
Classical Aesthetics→Expressive Aesthetics	.63	.74	n.s
Relationship	**MM**	**MF**	***p***
Classical Aesthetics→Credibility	.46	**.67**	.05
Expressive Aesthetics→Credibility	.15	.12	n.s
Classical Aesthetics→Expressive Aesthetics	.65	.65	n.s
Relationship	**FF**	**FM**	***p***
Classical Aesthetics→Credibility	.56	.82	n.s
Expressive Aesthetics→Credibility	.20	.00	n.s
Classical Aesthetics→Expressive Aesthetics	.75	.79	n.s
Relationship	**MM**	**FF**	***p***
Classical Aesthetics→Credibility	.46	.56	n.s
Expressive Aesthetics→Credibility	.15	.20	n.s
Classical Aesthetics→Expressive Aesthetics	.65	.75	n.s
Relationship	**MF**	**FM**	***p***
Classical Aesthetics→Credibility	.67	.82	n.s
Expressive Aesthetics→Credibility	.12	.00	n.s
Classical Aesthetics→Expressive Aesthetics	.65	**.79**	.05

The MGA for male vs. female observers shows that that there is no significant difference between both genders of observers in their evaluation of the behavior models. However, at the level of personalization, the MGA for MM vs. MF shows that there is a significant difference between both groups in the relationship between *classical aesthetics* and *perceived credibility* ($p < 0.05$). Moreover, there is a significant difference between MF and FM with respect to the relationship between *classical aesthetics* and *expressive aesthetics* ($p < 0.05$).

5 DISCUSSION

We presented a path model to investigate: (1) which of the two dimensions of *visual aesthetics* has a stronger influence on the *perceived credibility* of the visual design of behavior models aimed at changing behavior; and (2) the moderating effect of gender-based personalization. In this section, we discuss the validations of two of our main hypotheses and their implications, our contributions and the limitations of our study.

5.1 Validation of Hypotheses

Our path analysis (Figures 3 and 4) shows that, regardless of the gender of the observer, two of our three hypotheses are supported: *classical aesthetics* has a significant influence on *perceived credibility* (H1); (2) *classical aesthetics* has a significant influence on *expressive aesthetics* (H3); but (3) *expressive aesthetics* has no significant influence on *perceived credibility* (H2). The validation of H1 suggests that users rely on the *classical* dimension—and not the *expressive* dimension—of *visual aesthetics* in the evaluation of the *perceived credibility* of behavior model design. This finding replicates that of Oyibo et al. [6] in the tourism domain, in which the authors found that the *classical* dimension of *aesthetics* has a stronger influence than the *expressive* dimension on *perceived credibility*. The confirmation of H1 in the health domain suggests that *visual aesthetics* attributes, such as cleanness, clarity and pleasantness (encompassed in simplicity and orderliness), have more impact on the *perceived credibility* of the behavior model design than complex attributes (e.g., fascination, sophistication, creativity, etc.). This is well reflected in the following participant's comment: "*It's really clear, it is very credible because it's an actual exercise and it shows which muscles are working. It's not creative because it just shows a woman doing squats*" [P116, squat]. Despite commenting that the squat behavior model design is not creative (low in *expressive aesthetics*), participant 116 still considered the design "very credible" owing to its perceived clarity (*classical aesthetics*). Another comment that reflects participants' judgment of *credibility* based on mainly *classical aesthetics* is: "*The squat video is self explanatory and clean. Perfect for my needs*" [P96, squat].

A possible explanation for why *classical aesthetics* has a stronger influence on *perceived credibility* than *expressive aesthetics* does is that fitness apps modeling behavior are more of utilitarian than hedonic systems. Utilitarian systems provide instrumental value to the user. They *"address tasks and activities where user motivation in using the system is driven by the expectation of a reward or benefit external to his interaction with the system*" (p. 620) [21]. In the case of fitness apps modeling behavior, the user's expected (external) benefit is the favorable outcomes in engaging in physical activity, which could be physical or social. On the contrary, hedonic systems provide entertainment or enjoyment to the users. In other words, they satisfy the intrinsic needs of the user. This explains why *expressive aesthetics* (manifested in fascinating design and use of special effects [11]) does not influence *perceived credibility*. However, classical *aesthetics* (which is related to the traditional notion of *usability* such as orderliness, cleanness, clarity, etc. [11]) does influence *perceived credibility*, regardless of the gender of the observer and behavior model. Table 4 shows a cross-section of participants' comments, which reflect their concern with the *classical* dimension of *aesthetics* in the evaluation of fitness apps. For example, P212 commented that s/he is satisfied with the squat behavior model design because it is simple and clear (nothing distracting) and easy to follow.

Table 4: participants' comments on classical aesthetics

	Participants' Comment
1	"*The visual design is very simple and clear. The steps are easy to follow and the image of the person is a great illustration of how to properly do the squat. There's nothing distracting on the screen and the steps are simple*" [P212, Squat].
2	"*I like the visual design. It's uncluttered and simple. It does a good job demonstrating the exercise in an easy to follow way and I like that it also illustrates the muscle groups being used. Although the visuals are nothing over-the top or groundbreaking, they are effective at accomplishing their goal of demonstrating an exercise*" [P476, Push-Up].
3	"*The design is simple and i like that. The instructions are as well and the figure's movements are clear and easy to follow*" [P30, Squat].
4	"*I think that it is cleanly and clearly designed. I think that it is easy to follow and understand*" [P403, Push-Up].
5	"*I really like the design of the app, I think it's clear and easy to follow…"* [P263, Squat].

5.2 Findings based on Multigroup Analysis

With respect to the personalization of the gender of the behavior models to the user (Table 3), the significant difference we found between the group that evaluated behavior models of the same gender and the group that evaluated behavior models of the opposite gender is with respect to MM vs. MF. The effect of *classical aesthetics* on *perceived credibility* is significantly stronger for the opposite-gender MF dyad ($\beta = 0.67$, $p < 0.001$) than for the same-gender MM dyad ($\beta = 0.46$, $p < 0.001$). This finding seems to be replicated in the MGA for FF ($\beta = 0.56$, $p < 0.001$) vs. FM ($\beta = 0.82$, $p < 0.001$)—only that the numerical difference between both groups is not significant. Despite the difference between FF and FM is not statistically significant, the significant difference between MM and FM ***seems to suggest that users rely on classical aesthetics more in the judgment of credibility of the behavior model design when the behavior models are contra-tailored. In contrast, users seem to rely on expressive aesthetics more***

in the judgment of credibility of the behavior model design when the behavior models are personalized. For example, with respect to MM vs MF, we see that the effect of *classical aesthetics* on *perceived credibility* is significantly stronger for MF ($\beta = 0.67$, $p < 0.001$) than for MM ($\beta = 0.46$, $p < 0.001$), while the effect of *expressive aesthetics* on *perceived credibility* is non-significantly stronger for MM ($\beta = 0.15$, p = n.s) than for MF ($\beta = 0.12$, p = n.s). Similarly, we see that the effect of *classical aesthetics* on *perceived credibility* is non-significantly stronger for FM ($\beta = 0.82$, $p < 0.001$) than FF ($\beta = 0.56$, $p < 0.001$), while the effect of *expressive aesthetics* on *perceived credibility* is non-significantly stronger for FF ($\beta = 0.20$, p = n.s) than MF ($\beta = 0.00$, p = n.s). However, given the non-replication of the significant difference in the *classical aesthetics-perceived credibility* relationship between MM and MF with that between FF and FM, there is a need for further research into the above highlighted proposition.

5.3 Main Findings and Contributions

Based on the significant relationship between *classical aesthetics* and *perceived credibility* ($p < 0.001$), we recommend the following general design guideline:

> *When designing persuasive systems with utilitarian value or benefit, such as fitness applications featuring behavior models (virtual coach) aimed at changing behavior, HCI designers should focus on enhancing classical aesthetics (cleanness, orderliness, clarity and simplicity) to enhance the perceived credibility of the fitness applications.*

Our main contribution to the body of knowledge is that we replicated the above prior finding in the tourism domain [6] in the health domain, indicating that the hypothesis—*classical aesthetics has a stronger effect than expressive aesthetics on the perceived credibility of utilitarian systems*—is likely to cut across domains. Aside from the relationship between *classical aesthetics* and *perceived credibility*, we showed that there is a significant relationship between the two dimensions of *visual aesthetics*, which confirms our third hypothesis (H3). This suggests that the higher users' perception of a fitness app to be *classically aesthetic*, the higher will be their perception of the app as *expressively aesthetic*. However, a high perception of the app as *expressively aesthetic* does not result in a high perception of the app as *credible*, indicating that *expressive aesthetics* does not play a mediating role between the effect of *classical aesthetics* on *perceived credibility*.

5.4 Limitations and Future Work

Our study has some limitations. The main limitation is that our findings are based on users' perception (i.e., self-report). As such, we cannot guarantee that they will generalize to an actual context in which the users have to use the fitness app and answer the posed questions thereafter. The second limitation of our study is that most of the participants are from Canada and the United States. This may threaten the generalization of our findings to other demographics. Thus, in future work, we look forward to investigating our findings among other demographics as well as by using an actual fitness app.

6 CONCLUSION

This paper presents the results of an empirical study among 669 participants to investigate which of the two dimensions of *visual aesthetics* (*classical* and *expressive*) has a stronger impact on the *perceived credibility* of exercise behavior model design in fitness apps. Our path analysis shows that, regardless of the gender-based personalization of the behavior model to the user, *classical aesthetics* has a stronger influence than *expressive aesthetics* on *perceived credibility*. Our finding underscores the need for designers of persuasive systems with utilitarian benefits, such as fitness apps, to emphasize *classical aesthetics* (cleanness, orderliness, clarity and simplicity) more than *expressive aesthetics* (enrichment and complexity). This will go a long way in enhancing the *perceived credibility* of health apps aimed at changing behaviors.

REFERENCES

[1] B. J. Fogg, *Persuasive Technology: Using Computers to Change What We Think and Do*. Morgan Kaufmann, 2003.

[2] H. Oinas-kukkonen and M. Harjumaa, "Persuasive Systems Design Key Issues, Process Model, and System Features," *Commun. Assoc. Inf. Syst.*, vol. 24, no. 28, pp. 485–500, 2009.

[3] B. J. Fogg, C. Soohoo, D. Danielson, L. Marable, J. Stanford, and E. R. Tauber, "How Do People Evaluate a Web Site ' s Credibility ? Results from a Large Study," 2002.

[4] K. Oyibo and J. Vassileva, "The Interplay of Aesthetics, Usability and Credibility in Mobile Website Design and the Moderation Effect of Gender," *SBC J. Interact. Syst.*, vol. 7, no. 1, 2017.

[5] K. Oyibo and J. Vassileva, "The Interplay of Aesthetics, Usability and Credibility in Mobile Websites and the Moderation by Culture," 2016.

[6] K. Oyibo, A. Ifeoma, R. Orji, and J. Vassileva, "What Drives the Perceived Credibility of Mobile Websites: Classical or Expressive Aesthetics?," in *20th Int'l Conference on Human-Computer Interaction*, 2018, pp. 576–594.

[7] D. Robins and J. Holmes, "Aesthetics and credibility in web site design," *Inf. Process. Manag.*, vol. 44, no. 1, pp. 386–399, 2008.

[8] C. Abraham and S. Michie, "A taxonomy of behavior change techniques used in interventions.," *Health Psychol.*, vol. 27, no. 3, pp. 379–87, 2008.

[9] M. Moshagen and M. T. Thielsch, "Facets of visual aesthetics," *Int. J. Hum. Comput. Stud.*, vol. 68, no. 10, pp. 689–709, 2010.

[10] N. Tractinsky, "Visual Aesthetics," *Interaction-Design*, pp. 1–71, 2002.

[11] T. Lavie and N. Tractinsky, "Assessing dimensions of perceived visual aesthetics of web sites," *Int. J. Hum. Comput. Stud.*, vol. 60, no. 3, pp. 269–298, 2004.

[12] 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.

[13] B. J. Fogg and H. Tseng, "The elements of computer credibility," in *Proceedings of the SIGCHI conference on Human factors in computing systems the CHI is the limit - CHI '99*, 1999, pp. 80–87.

[14] F. Alsudani and M. Casey, "The Effect of Aesthetics on Web Credibility," *23rd Br. HCI Gr. Annu. Conf. People Comput. Celebr. People Technol.*, pp. 512–519, 2009.

[15] 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.

[16] K. Oyibo, R. Orji, I. Adaji, O. Babatunde, M. Azizi, and J. Vassileva, "Persuasive Effect of Behavior Model Design in Fitness Apps," in *Proceedings of the 2018 Conference on User Modeling Adaptation and Personalization - UMAP '18*, 2018, pp. 219–228.

[17] P. van Schaik and J. Ling, "The role of context in perceptions of the aesthetics of web pages over time," *Int. J. Hum. Comput. Stud.*, vol. 67, no. 1, pp. 79–89, 2009.

[18] L. Bergkvist and J. R. Rossiter, "The Predictive Validity of Multiple-Item Versus Single-Item Measures of the Same Constructs," *J. Mark. Res.*, vol. 44, no. 2, pp. 175–184, 2007.

[19] G. Sanchez, "PLS Path Modeling with R," *Berkley Trowchez Ed.*, 2013.

[20] K. K. Wong, "Partial Least Squares Structural Equation Modeling (PLS-SEM) Techniques Using SmartPLS," *Mark. Bull.*, vol. 24, no. 1, pp. 1–32, 2013.

[21] H. Wigelius and H. Väätäjä, "Dimensions of context affecting user experience in mobile work," in *IFIP Conference on Human-Computer Interaction*, 2009, pp. 604–617.