Demographic factors have little effect on aesthetic perceptions of icons: a study of mobile game icons

Henrietta Jylhä and Juho Hamari
Gamification Group, Tampere University, Tampere, Finland

Received 7 July 2020
Revised 7 October 2020
11 January 2021
22 June 2021
Accepted 22 June 2021

Abstract

Purpose – Customization by segmenting within human–computer interaction is an emerging phenomenon. Appealing graphical elements that cater to user needs are considered progressively important, as the way a graphic is visually represented can greatly contribute to the interaction. However, aesthetic perceptions are subjective and may differ by target group. Understanding variations in user perceptions may aid in design processes; therefore, we set out to investigate the effects of demographic differences relating to perceptions of graphical user interface (GUI) element (i.e. game app icon) aesthetics.

Design/methodology/approach – The authors employed a vignette experiment with random participant (n = 513) assignment to evaluate 4 icons from a total of 68 pre-selected mobile game icons using semantic differential scales. This resulted in a total of 2052 individual icon evaluations. Regression analyses were performed with the effects of age, gender and time using graphical user interfaces (i.e. app stores) and the interactions of these variables relating to perceptions of GUI element aesthetics.

Findings – The results indicate that, overall, demographic factors have relatively little effect on how icons are perceived. Significant relations suggest that experienced users, younger audiences and women are more critical in their perception of aesthetic excellence, and that perceptions change for younger women. The implications of the findings are discussed via adaptive decision-making theory.

Originality/value – In the context of graphical user interface element aesthetics, demographic differences have received minimal attention as moderating variables regardless of their relevance in design and development. Hence, it merits further research.

Keywords Iconography, Aesthetics, Demographics, User perception, Graphical user interface, Human–computer interaction

Paper type Research paper

1. Introduction

Demographic differences in designing aesthetically pleasing graphical user interface (GUI) elements have become prevalent due to increasing demands for customization within human–computer interaction (Norman, 2004; Tractinsky *et al.*, 2000). As a wide variety of daily communication is realized via user interfaces of different devices, designers are presented with new opportunities and challenges to create visually effective GUI elements for their targeted consumer group. Moreover, perceptions of successful (i.e. appealing) visual aesthetics are subjective (Zen and Vanderdonckt, 2016), which complicates creating balanced user experiences for critical masses. Especially in mobile environments, the adoption of mobile game applications is a complex entity of varying perceptions, such as gender, content price and quality and time spent playing mobile games (Pappas *et al.*, 2019). Therefore, insight into what aspects of GUI element aesthetics are preferred by segmentation is needed.

User interfaces that adapt to individual preferences have been shown to lead to higher ratings in look and feel as well as long-term usage of platforms (Debevc *et al.*, 1996; Hartmann

Academy of Finland Flagship Programme [337653 - Forest-Human-Machine Interplay (UNITE)].

Internet Research
Vol. 32 No. 7, 2022
pp. 87-110
Emerald Publishing Limited
1066-2243
DOI 10.1108/INTR-07-2020-0368

et al., 2007a; Sarsam and Al-Samarraie, 2018). Considering that ineffective interface usability tends to affect older age groups (Johnson and Finn, 2017) due to visual acuity changes (Huang, 2013), and that age is likely to contribute to users' skill level and experience with technology (KnowItAll Ninja, 2016), it can be considered a meaningful factor in GUI aesthetics and design which merits for further research on the topic.

Regarding gender differences in the field of human-computer interaction and visual aesthetics, the norm has been that preferences of male and female users differ to a significant degree (Genuine, 2013); however, new trends of more unisex patterns have been discovered (Morris *et al.*, 2005). In the future, offering gender-neutral options for user interfaces could be one solution to the possible minimizing of gender differences (Boiano *et al.*, 2006). Due to the change of the cultural atmosphere, there is a need to examine the effects of gender in this context.

Time interacting with interfaces contributes to impressions on aesthetics, nevertheless, this topic has received relatively little attention especially considering mobile interfaces (Miniukovich and De Angeli, 2014). Time affects several user attributes, preferences and expectations (Hartmann *et al.*, 2008; Thüring and Mahlke, 2007) that can lead to various outcomes concerning interface design. The norms of device interaction suggest that, grave alterations to GUI designs may hinder user adjustment and lead to frustration, and thus gradual changes are advised (KnowItAll Ninja, 2016). As the frequency of use is related to aesthetic perceptions on a general level, it is an important variable in determining the subjective experience.

Prior research has indicated that not only the main effects of age, gender and time are to be investigated, but also the interactions of these demographics should be taken into account, as significant relationships have been found between, e.g. age and gender on technology adoption (Morris *et al.*, 2005) as well as gender and time on mobile entertainment (Hsiao and Chen, 2016; Pappas *et al.*, 2019). Research regarding demographic differences in relation to aesthetic perceptions of GUI elements is scarce at present. The rapid progress of GUI design further justifies the current undertaking.

As described, different results exist on user interface aesthetics and the trends regarding age, gender and time spent interacting with devices, thus more work is needed to understand how the interplays of these particular demographics may offer a deeper understanding on perceptions of GUI aesthetics, and how they may affect further design and research processes. To address this gap, we observe user perceptions on GUI aesthetics based on adaptive decision-making theory (Payne *et al.*, 1993). Used in previous evaluations of interface quality (e.g. Hartmann *et al.*, 2007a, b, 2008), this approach allows interpreting the results with a conception that user judgment is adaptive and based on the task, context and background-experience. This theory is valid particularly in choice situations where no single alternative is best on all attributes (Beresford and Sloper, 2008).

The large-scale quantitative demographic data in this study was collected via a vignette experiment with random participant ($n = 513$) assignment, where the task was to evaluate 4 icons from a total of 68 pre-selected game app icons across 4 categories (concrete, abstract, character and text) using semantic scales. This resulted in a total of 2052 individual icon evaluations. Based on the results, our study presents insight into the effects of age, gender and time using graphical user interfaces (i.e. *app stores*) and the interactions of these variables relating to perceptions of GUI element aesthetics. Knowledge of these relations allows for theoretical and practical guidelines in the design process of personalized graphical user interface elements.

2. Background

2.1 Aesthetics perceptions in graphical user interfaces

Visual aesthetics in graphical user interface design can be defined as aesthetically pleasing or attractive computer-based environments, reflecting the format in which the content and services are presented as well as the design look and feel and overall experience with a system (Ahmed *et al.*, 2009; Hartmann *et al.*, 2007b; Jennings, 2000). As a research field, it focuses on the

user's subjective judgment on how aesthetic a system or a product is (Lee and Koubek, 2011), an increasingly important area in human–computer interaction due to the wide adaptation of devices for everyday actions. Aesthetics within human–computer interaction can be divided into classical and expressive aesthetics (Ahmed *et al.*, 2009; Hartmann *et al.*, 2008; Lavie and Tractinsky, 2004). Classical aesthetics refers to clear designs, whereas expressive aesthetics refer to more creative designs. Especially concerning interface icons, aesthetic appeal has been described as mild aesthetic experiences that refer to the power to attract users (McDougall *et al.*, 2016). In system design, the structure of information has been linked with perceived aesthetics as well as usability (Ahmed *et al.*, 2009; Cyr, 2009). Interaction with user interfaces is realized via graphical elements providing intuitiveness and immediate visual feedback, such as windows, menus and icons (Linux Information Project, 2004). Aesthetics in graphical user interface design has been proven an integral part of a positive user experience as well as user engagement (Kurosu and Kashimura, 1995; Ngo *et al.*, 2000; Overby and Sabyasachi, 2014; Salimun *et al.*, 2010; Tractinsky *et al.*, 2000). Positive user experience is important for successful human-computer interaction, as the user may abandon an interface that is related with a negative experience. User experience is connected to visual aesthetics to an increasing extent (Debevc *et al.*, 1996; Hartmann *et al.*, 2007a; Sarsam and Al-Samarraie, 2018); hence, an attractive user interface is important when aiming for successful human-computer interaction as well as positive commercial performance (Gait, 1985; Lin and Yeh, 2010).

Perceptions of effective visual aesthetics have been attempted to assess via various theories and tools (e.g. Choi and Lee, 2012; Hassenzahl *et al.*, 2003; Maity *et al.*, 2015; Ngo *et al.*, 2000; Ngo, 2001; Ngo *et al.*, 2003; Salimun *et al.*, 2010; Zen and Vanderdonckt, 2016), yet robust guidelines for designing GUI elements are lacking due to the complexity of the topic. Prior research (Maity *et al.*, 2015; Ngo *et al.*, 2000) has found correlations between metric-based aesthetic value and the aesthetics ratings of design experts, artists and users. However, these results were only partly supported by a similar study (Zen and Vanderdonckt, 2016). Another study (Salimun *et al.*, 2010) contrasted prior literature (Ngo, 2001; Ngo *et al.*, 2003) in that some metrics, such as symmetry and cohesion, influence results more than others. In addition to metric-based instruments, aesthetic value of graphical user interfaces has been measured by survey-based methods (Choi and Lee, 2012; Hassenzahl *et al.*, 2003; Jylhä and Hamari, 2020) aligned with user perceptions. Prior contradictory results in evaluation theories and tools of aesthetics in interface research are perhaps due to analyzing user interfaces as entities (Zen and Vanderdonckt, 2016). As user interfaces essentially consist of several elements with different purposes, it motivates investigating GUI elements separately rather than as an entity. Therefore, in this study, we scaled the sample into single interface components, i.e., icons. While icons do not constitute a GUI solitarily, icon-based interfaces are highly common at present. This justifies using icons as study material for evaluating the effect of demographic differences within user perceptions of GUI element aesthetics.

2.2 Demographic differences in interaction design

Prior literature on the effects of demographic differences in human-computer interaction and aesthetics suggests an impact on user perceptions, motivations and design processes (Creusen, 2010; Johnson and Finn, 2017; Oyibo *et al.*, 2016, 2018). However, in the context of graphical user interface element aesthetics, demographic differences have received minimal attention as moderating variables (Oyibo *et al.*, 2018). Advancing knowledge in the topic is beneficial to scholars and practitioners alike as contributions of this study may be adapted in further examining and designing user interface systems within the context of human–computer interaction.

Regarding interface design and age, it has been indicated that younger people tend to focus more on hedonic pleasure, whereas older people prefer a more utilitarian approach (Hsieh *et al.*, 2004; Johnson and Finn, 2017; Wallendorf and Arnould, 1988). Research has

shown that younger people are more critical towards aesthetics than older people, who were found to be indifferent about color schemes, while younger people were found to prefer moderate-temperature (green and orange) to extreme temperature (blue and red) color schemes (Oyibo *et al.*, 2018). Thus, interface designers are prompted to put effort in aesthetics considerations in order to appeal to younger audiences. Prior studies regarding age and technology have indicated a digital divide between generations in which younger age groups are less affected by social influence due to early technology adoption (Morris and Venkatesh, 2000; Venkatesh *et al.*, 2000). Moreover, older generations tend to experience more anxiety relating to human-computer interaction than younger people, as the process of adapting to new devices may be more time-consuming with age due to the diminishing of cognitive abilities such as memory capacity, symbol and language comprehension (Chung *et al.*, 2010; Creusen, 2010; Johnson and Finn, 2017; Rousseau *et al.*, 1998). Therefore, prior literature (Johnson and Finn, 2017) has suggested a number of design guidelines (e.g. the use of large fonts, maintaining visual consistency) in order to accommodate the aging population. Differences in perceptions between young and old users and the digital divide between age groups in today's society motivates for further observation of age in this context.

Concerning the effects of gender, implications have been made in terms of decision-making and information processing in that male users concentrate more on pragmatic aspects of technology, while female users are driven by social motivators (Sun and Zhang, 2006; Venkatesh and Morris, 2000). This means that in general, men are more orientated towards completing tasks and achievements than women (Hoffman, 1972; Minton and Schneider, 1980). On the other hand, women are more concerned with influential motivators and have been considered to be less likely to enjoy the use of information technology (Creusen, 2010; Hoffman, 1972; Venkatesh and Morris, 2000). Relating to interface design, males tend to prefer functional aspects (i.e. usability and symmetry), while females prefer expressive aspects (i.e. beauty and emotional value) (Creusen, 2010; Henry, 2002; Oyibo and Vassileva, 2017; Tuch *et al.*, 2010; Wallendorf and Arnould, 1988). In this regard, men can be considered more instrumental, whereas women are more symbolic and concerned with appearance. Furthermore, females have been found to be more sensitive to color and visual complexity in the context of user interfaces than males (Creusen, 2010; Reinecke and Gajos, 2014; Smith, 1995). A study in the context of mobile service adoption found no gender differences (Leong *et al.*, 2013). However, this has been countered by discovering that male and female users of mobile systems have different motivations, for example, males favor status and value and females prefer social and utilitarian orientations (Liu and Guo, 2017). This raises the need to examine the effect of gender especially in mobile environments.

In addition to age and gender, the effects of time using graphical interfaces (i.e. *app stores*) is to be taken into account. Prior literature has indicated that time affects user attributes, such as knowledge and skill level, as well as perceptions of system features, such as design and functionality (Lee and Koubek, 2011; Thüring and Mahlke, 2007). Frequent use of devices has shown to affect user preferences and expectations of visual aesthetics (Lee and Koubek, 2011). Moreover, users have been found to be selective with aesthetics based on experience (Hartmann *et al.*, 2008). Prior research on mobile entertainment has identified that time spent interacting with mobile systems affect user intentions and motivations, such as mobile game preference (Hsiao and Chen, 2016) and the level of investment on downloading mobile games (Pappas *et al.*, 2019). Involvement with GUI elements may impact users in several ways in regards to skill level, user experience, decision-making processes and perceptions of aesthetics. However, this topic has received relatively little attention especially considering mobile interfaces (Miniukovich and De Angeli, 2014). On the basis of prior literature and due to the lack of recent research, time is considered a valid factor in this study.

Demographic effects in the context of technology have shown to form multiple configurations of causal conditions. Prior research on technology adoption (Morris *et al.*, 2005)

has found a trend for more unisex pattern among younger people, suggesting that both younger women and men have received greater exposure to technology compared with the older generation, thus minimizing gender differences in this area. Furthermore, prior literature (Pappas *et al.*, 2019) has found a link between gender, content price and quality, as well as time spent playing mobile games: females who spend a lot of time playing games are more willing to overspend if the content is of high quality. This justifies the interpretation of effects between variables with respect to each other. Therefore, in addition to the independent variables age, gender and time, we employ the interaction effects of these variables. Interactions assess the relationship between an independent variable and dependent variable, moderated by a third variable (Aiken and West, 1991). This indicates that a third variable might influence the relationship between an independent and dependent variable, allowing for the observation of a more complex model where not only the main effects are studied. This can greatly expand understanding the relationships among variables in the model (Sweet and Grace-Martin, 2011).

2.3 Aesthetics, demographics and adaptive decision-making theory

In this paper, we observe user perceptions of GUI aesthetics in a theoretical framework of adaptive decision-making. The adaptive decision-making theory posits that an individual's use of decision strategies is an adaptive response of a limited-capacity information processor to the demands of complex decision tasks (Payne *et al.*, 1993). A person's repertoire of decision-making strategies depends on many factors, such as cognitive development, experience, and more formal training and education (Hartmann *et al.*, 2007a). As prior knowledge determines which strategies are available to a decision-maker, our elaboration of this theory hypothesizes that age, gender and experience are amongst the variables that affect decision-making behavior, contributing to individual differences.

The handful of studies that have investigated the relationship between interface aesthetics and demographics from the perspective of adaptive decision-making theory have consensus that the user's background plays an important role in the judgment of aesthetic appeal (Hartmann *et al.*, 2007a, b, 2008). These studies on aesthetics, usability and content relating to user interfaces theorize that preferences for user interface designs when the scenario of use is critical will be based on in-depth consideration, whereas for less serious scenarios, preferences will be based on selecting designs by general aesthetic impressions. The studies conclude that design priorities for aesthetics should be matched to user profile. Designers should not only know their audience, but also the audience's decision-making habits (i.e. preferences and expectations) that depend on interactions between decision-making criteria (e.g. design qualities such as content, aesthetics, functionality, usability).

As literature on this topic is limited, further investigation is justified. In the milieu of this theoretical framework and the study experiment, we expect users to evaluate the icon material by the strategy of desirability to the decision maker, a trait that is likely to be affected by demographic factors and background-experience. In particular, we hypothesize that the pattern of women being more drawn to the expressive aspects than men will continue, and that the time interacting with devices will have an effect on perceptions of visual aesthetics, i.e., the more time is spent, the more critical the users are towards the design aspects.

3. Methods and data

3.1 Participants

A nonprobability convenience sample was composed initially of 569 respondents who each assessed 4 game app icons through a survey-based within-subjects vignette experiment.

A within-subjects approach was chosen as opposed to between-subjects approach in order to expose each participant to all conditions (i.e. 4 icon evaluations by category) of the experiment. Due to insufficient representation, 15 responses without identifiable gender were removed. Additionally, 41 responses from older age groups were identified as outliers and removed, resulting in a total of 513 respondents with 2052 icon evaluations. Please refer to Table 1 for demographic details of participants.

The experiment was a self-administered online task. The aim was to gather data by exposing the participants close to a realistic setting outside an authentic app store context. The majority of participants resided in Finland (93.0%). The gender split across participants was rather equal, as only slightly more than half were male (52.4%). The mean age was 25.49 years (SD = 4.67 years; 16–39 years). As the majority of the respondents were from the same age group, the results of this study can be considered more representative of younger age groups. Most participants were university students (65.7%) and had a university-level education (41.1%). The majority of participants (40.2%) browsed app stores once per week. Most participants (75.6%) did not download any game apps on a weekly basis. Missing data (1.8%) was encountered for these two aforementioned items, as the frequency of app store usage and mobile game downloads were only asked from those who use a smartphone. To counter possible bias in the experiment, participants who did not download game apps frequently were instructed to answer based on their expectations of game app icons they

Table 1. Demographic information

			n	%
Gender	Male		269	52.4
	Female		244	47.6
Age by gender (SD = 4.67)	−19	Male	14	2.7
(Mean = 25.49)		Female	8	1.6
(Median = 25.00)		Total	22	4.3
	20–24	Male	105	20.5
		Female	120	23.4
		Total	225	43.9
	25–29	Male	103	20.1
		Female	75	14.6
		Total	178	34.7
	30–34	Male	27	5.3
		Female	26	5.0
		Total	53	10.3
	35–39	Male	20	3.9
		Female	15	2.9
		Total	35	6.8
Times browsing app stores	0		147	28.7
	1		206	40.2
	2		74	14.4
	3		34	6.6
	4		10	1.9
	5		12	2.3
	More than 5		21	4.1
	Missing		9	1.8
Game apps downloaded (per week)	0		388	75.6
	1		94	18.3
	2		13	2.5
	3		7	1.4
	4		2	0.4
	Missing		9	1.8

might interact with. Two participants were randomly chosen and awarded a prize (Polar Loop 2 Activity Tracker). No other participation fees were paid. Participants were informed of the purpose of the study and assured anonymity.

3.2 Materials

Sixty-eight game app icons from Google Play Store were selected for the study. The decision to narrow down the sample to game app icons was made to eliminate further variability that might stem from the nature of the app and thus increase internal validity of the experiment, but also external validity in terms of results applied to the game icons. In order to avoid any systematic bias, 4 icons corresponding to dominant icon styles (concrete, abstract, character and text) were selected from each of 17 categories for game apps (action, adventure, arcade, board, card, casino, casual, educational, music, puzzle, racing, role playing, simulation, sports, strategy, trivia and word). Because icon design for app stores is category-dependent (Shu and Lin, 2014), we considered it justified to include icons from all categories. Prior literature highlights the relevance of concreteness and abstractness in icon design (e.g. Arend *et al.*, 1987; Blankenberger and Hahn, 1991; Dewar, 1999; Hou and Ho, 2013; Isherwood *et al.*, 2007; McDougall and Reppa, 2008; McDougall *et al.*, 1999; McDougall *et al.*, 2000; Moyes and Jordan, 1993; Rogers and Oborne, 1987); hence, they were included in this experiment. Looking at the icons on app stores, characters and typography are prevalent elements usually seen on app icons. It has been argued that faces on app icons are widely used because of the immediate impact and memorability they have due to neural processing of facial expressions (Chartboost, 2015). Furthermore, as the study design is based on prior research (Shaikh, 2009) on onscreen typeface and usage, text elements were included. During the selection phase we ensured that one icon from each category was dominantly characteristic of one of these 4 attributes.

Additional criteria were the publishing date of the apps and the number of installs and reviews they had received at the time of selection. Since the icons in the experiment were chosen during December 2016, the acceptable publishing date for the apps was determined to range from December 3–17 2016. No more than 500 installs and 30 reviews were permitted. The aim of this was to choose new app icons to eliminate the chance of app and icon familiarity and thus, systematic bias. Moreover, the goal was to have as visually rich a sample of icons as possible, meaning that several different computer graphic techniques were included, such as 2D and 3D rendered images. The icons are presented in Table 2.

3.3 Measurements

Semantic differential scale was used to measure respondent evaluations of aesthetic aspects of the icons. A total of 22 adjective pairs was formulated and assigned to each icon. The polarity of the adjective pairs was reversed so that perceivably positive and negative adjectives did not align on the same side of the scale. All of the adjective pairs were chosen according to prior research (Shaikh, 2009) on onscreen typeface design and usage. Additionally, adjectives related to icons were added as suggested per previous literature on effective icon design. These adjectives include concrete and abstract (Arend *et al.*, 1987; Blankenberger and Hahn, 1991; Dewar, 1999; Hou and Ho, 2013; Isherwood *et al.*, 2007; McDougall and Reppa, 2008; McDougall *et al.*, 1999, 2000; Moyes and Jordan, 1993; Rogers and Oborne, 1987), simple and complex (Choi and Lee, 2012; Goonetilleke *et al.*, 2001; McDougall and Reppa, 2008; McDougall and Reppa, 2013; McDougall *et al.*, 2016) as well as unique and ordinary (Creusen and Schoormans, 2005; Creusen *et al.*, 2010; Dewar, 1999; Goonetilleke *et al.*, 2001; Huang *et al.*, 2002; Salman *et al.*, 2010). Furthermore, adjective pairs that were added to specifically measure the aesthetics of the icons include professional and unprofessional, colorful and colorless, realistic and unrealistic as well as two-dimensional and three-dimensional (Jylhä and Hamari, 2019).

Category	Concrete	Abstract	Character	Text
Action				
Adventure				
Arcade				
Board				
Card				
Casino				
Casual				
Educational				
Music				
Puzzle				
Racing				
Role Playing				
Simulation				
Sports				
Strategy				
Trivia				
Word				

Table 2.
Icons in the study

Developed further into a five-factor model entitled VISQUAL (Jylhä and Hamari, 2020), an instrument for measuring visual qualities of graphical user interface elements, the scale was used to observe underlying latent constructs in this study. VISQUAL consists of the aforementioned adjective pairs that were further divided into the following dimensions: *Excellence/Inferiority*, *Graciousness/Harshness*, *Idleness/Liveliness*, *Normalness/Bizarreness* and *Complexity/Simplicity*.

Table 3 lists the VISQUAL constructs and adjective pairs. Two versions of the model exist, the initial model with 22 adjective pairs and an adjusted model of 15 adjective pairs. In Table 3, the bolded adjective pairs represent those included in the adjusted model of 15 adjective pairs. Table 3 also presents an overview of the means and standard deviations. There were no outlier values and the range between the lowest and highest scores clustered closely to the average even though the 68 icons were quite different from each other. All the mean scores were between 3.5 and 4.5 for each evaluation. This indicates little skewness in the data.

Additional to the semantic scales, a seven-point Likert scale was utilized to measure the degree of disagree-agreement of the respondents with respect to the likelihood of them clicking, downloading, and purchasing the imagined app behind the icon with an instruction title: "Overall evaluation (judging by the icon alone)" followed by questions: "Compared to the mobile game icons I usually click, I would click this icon," "Compared to the icons of mobile games I usually download, I would click this icon" and "Compared to the icons of mobile games I usually purchase, I would click this icon." Respondents were provided the following options on the seven-point scale: "Strongly disagree," "Disagree," "Somewhat disagree," "Neither agree nor disagree," "Somewhat agree," "Agree" and "Strongly agree." Moreover, respondents were asked to give an overall evaluation score for the design of each icon by grading them on a seven-point scale to further assess consumer perceptions of icon successfulness.

Factor	Adjective pair	Mean	SD
Excellence/	*Good–Bad*	4.34	1.641
Inferiority	*Professional–Unprofessional*	4.22	1.736
	Beautiful–Ugly	4.57	1.618
	Expensive–Cheap	4.83	1.563
	Strong–Weak	3.93	1.464
Graciousness/	Soft–Hard	3.81	1.545
Harshness	*Relaxed–Stiff*	4.47	1.560
	Masculine–Feminine	4.34	1.388
	Delicate–Rugged	4.42	1.368
	Happy–Sad	3.80	1.507
	Colorful–Colorless	3.77	1.810
	Warm–Cool	3.97	1.436
Idleness/	*Fast–Slow*	3.87	1.576
Liveliness	*Quiet–Loud*	4.12	1.601
	Exciting–Calm	3.96	1.452
	Active–Passive	3.97	1.708
	Young–Old	3.98	1.611
Normalness	*Concrete–Abstract*	4.03	1.998
Bizarreness	*Realistic–Unrealistic*	4.22	1.592
	Ordinary–Unique	4.60	1.651
Complexity/	*Simple–Complex*	4.69	1.669
Simplicity	*Three-dimensional–Two-dimensional*	3.33	1.863

Table 3. Constructs in VISQUAL, means and standard deviations (adjusted 15 model items italics)

3.4 Procedure

The data was collected through a survey-based vignette experiment. Respondents were provided the purpose of the study after which they were guided to fill out the survey. The survey consisted of three or four parts depending on the choice of response. The first part mapped out mobile game and smartphone usage with the following questions: "Do you like to play mobile games?", "In an average day, how much time do you spend playing mobile games?" and "How many smartphones are you currently using?". The second part included more specific questions about the aforementioned, e.g., the operating system of the smartphone(s) in use, the average number of times browsing app stores per week and the amount of money spent on app stores during the past year, as well as the importance of icon aesthetics when interacting with app icons. If the respondent answered that they do not use a smartphone in the first part, they were assigned directly to the third part.

In the third part, the respondent was asked to evaluate app icons using seven-point semantic differential scales. Prior to this, the following instructions were given on how to evaluate the icons: "In the following section you are shown pictures of four (4) mobile game icons. The pictures are shown one by one. Please evaluate the appearance of each icon according to the adjective pairs shown below the icon. In each adjective pair, the closer you choose to the left or right adjective, the better you think it fits to the adjective. If you choose the middle space, you think both adjectives fit equally well." The respondent was reminded that there are no right or wrong answers and was then instructed to click "Next" to begin. The respondent was shown one icon at a time and was asked to rate the 22 adjective pairs under the icon graphic with an initial "In my opinion, this icon is. . .." Each respondent was randomly assigned four icons to evaluate, one from each category of pre-selected icon attributes (abstract, concrete, character and text). After the semantic scales, the participant rated their willingness to click the icon as well as download and purchase the imagined app that the icon belongs to, by using a seven-point Likert scale on the same page with the icon. Last, demographic information (age, gender, etc.) was asked. The survey took about 10 minutes to complete. The survey was implemented via Surveygizmo, an online survey tool. All content was in English. The data was analyzed with IBM SPSS Statistics version 24 and Microsoft Office Excel 2016. The following section describes the results of the analysis.

4. Results

Regression analyses on the dependent variables (VISQUAL models and individual adjective pairs) were performed with age, gender and times browsing app stores (per week), as well as with the interaction terms of independent variables, namely age × gender, age × time, gender × time and age × gender × time. In the analyses, the ratio-scale variable of age and time was used instead of the ordinal scales in Table 1. The independent variables age and time were centered prior to the analyses (Aiken and West, 1991), and the interaction terms were created from the prior centered variables. Prior to the analyses, multicollinearity test was performed on the independent variables as well as the interaction terms with variance inflation factors (VIF). No critical levels of multicollinearity were found between the variables. The polarity of the adjective pairs was rotated so that perceivably positive and negative adjectives did not align on the same side of the scale. Prior to the analyses, items were reverse coded as necessary. First, regression analyses according to the VISQUAL model with 15 adjective pairs were performed. Please refer to Tables 4 and 5 for regression results.

When examining the results for statistically significant effects concerning age, gender and time, the results indicate that age ($\beta = -0.152$, $p = 0.033$) affects the *Excellence/Inferiority* dimension.

When observing these results, some statistically significant interactions between the independent variables were found. A two-way interaction between age and gender was found

for the dimension *Excellence/Inferiority* ($\beta = 0.170, p = 0.17$). Concerning age and time, a two-way interaction was found for the dimension *Normalness/Bizarreness* ($\beta = 0.167, p = 0.37$). No significant effect was found between gender and time or age, gender and time.

Second, regression analyses on the dependent variables according to the VISQUAL model with 22 adjective pairs were performed. Refer to Tables 6 and 7 for regression results.

Here, the results indicate that age ($\beta = -0.170, p = 0.17$) and time browsing app stores ($\beta = 0.173, p = 0.41$) affect the *Excellence/Inferiority* dimension.

Similar to the previous regression analyses (Table 5), a two-way interaction between age and gender was found for the dimension *Excellence/Inferiority* ($\beta = 0.193, p = 0.007$). Concerning age and time, a two-way interaction was found for the dimension *Normalness/Bizarreness*. ($\beta = 0.208, p = 0.009$). Additionally, a three-way interaction was found for the dimension *Normalness/Bizarreness* for age, gender and time ($\beta = -0.182, p = 0.025$). No significant effect was found between gender and time.

Lastly, regression analyses on the dependent variables as the individual 22 adjective pairs were performed. Refer to Tables 8 and 9 for regression results (organized per significant effects for further clarification).

When examining the results for statistically significant effects concerning age, gender and time, the adjective pairs *expensive–cheap* ($\beta = -0.163, p = 0.022$), *strong–weak* ($\beta = -0.172, p = 0.015$), *professional–unprofessional* ($\beta = -0.164, p = 0.022$) and *soft–hard* ($\beta = 0.157, p = 0.027$) were predicted by age. Regarding gender, *soft–hard* ($\beta = 0.058, p = 0.10$) and *warm–cool* ($\beta = 0.053, p = 0.019$) were predicted by gender. Finally, the adjective pair *ordinary–unique* ($\beta = -0.175, p = 0.039$) was affected by time spent browsing app stores on a weekly basis.

When examining these results, several statistically significant interactions between the independent variables were found. A two-way interaction between age and gender was found for the adjective pairs *expensive–cheap* ($\beta = 0.177, p = 0.013$), *strong–weak* ($\beta = 0.203, p = 0.004$), *professional–unprofessional* ($\beta = 0.181, p = 0.011$), *ordinary–unique* ($\beta = -0.156, p = 0.028$) and *good–bad* ($\beta = 0.147, p = 0.039$). Concerning age and time, a two-way interaction was found for the adjective pair *concrete–abstract* ($\beta = 0.170, p = 0.033$). A three-way interaction between age, gender and time was found for the adjective pairs *concrete–abstract* ($\beta = -0.166, p = 0.040$) and *exciting–calm* ($\beta = -0.162, p = 0.046$). No significant effect between gender and time was found.

5. Discussion

This study investigated the effects of age, gender and time using graphical user interfaces (i.e. *app stores*) relating to perceptions of GUI element (i.e. *game app icon*) aesthetics. The results

	Beta	*t*	*p*	Beta	*t*	*p*	Beta	*t*	*p*
		Age			Gender[a]			Time	
Excellence–Inferiority	*−0.152**	*−2.134*	*0.033*	0.008	0.346	0.729	0.159	1.877	0.061
Graciousness–Harshness	0.045	0.635	0.525	−0.011	−0.496	0.620	0.044	0.513	0.608
Idleness–Liveliness	−0.064	−0.899	0.369	−0.018	−0.790	0.430	0.075	0.883	0.377
Normalness–Bizarreness	−0.014	−0.202	0.840	−0.030	−1.331	0.183	0.112	1.321	0.187
Complexity–Simplicity	−0.020	−0.933	0.351	0.024	1.125	0.261	0.032	1.435	0.151

Note(s): * = $p < 0.05$, ** = $p < 0.01$, statistically significant effects italicized. [a]Females were coded with the higher variable value

Table 4. Regression analyses (15 items) adjusted model with age, gender and time spent browsing app stores (per week)

Table 5.
Regression analyses (15 items) with Age × Gender, Age × Time, Gender × Time, and Age × Gender × Time

	Beta	*t*	*p*	Beta	*t*	*p*	Beta	*t*	*p*	Beta	*t*	*p*
	Age × Gender			Age × Time			Gender × Time			Age × Gender × Time		
Excellence–Inferiority	*0.170**	*2.391*	*0.017*	0.024	0.302	0.762	−0.129	−1.508	0.132	0.006	0.077	0.938
Graciousness–Harshness	−0.015	−0.209	0.835	−0.035	−0.438	0.661	0.052	0.639	0.856	0.052	0.639	0.523
Idleness–Liveliness	0.080	1.121	0.262	0.138	1.733	0.083	−0.126	−1.551	0.186	−0.126	−1.551	0.121
Normalness–Bizarreness	0.017	0.239	0.811	*0.167**	*2.090*	*0.037*	−0.128	−1.488	0.137	−0.153	−1.890	0.059
Complexity–Simplicity	−0.033	−1.521	0.128	−0.023	−0.866	0.386	−0.034	−1.536	0.125	−0.018	−0.653	0.514

Note(s): * = $p \leq 0.05$, ** = $p < 0.01$, statistically significant effects italicized. [a]Females were coded with the higher variable value

indicate that, overall, demographic factors have relatively little effect on how icons are perceived.

Observing the effects concerning age, gender and time with the VISQUAL models as dependent variables (Tables 4 and 6), statistically significant effects were found for age and time within the Excellence/Inferiority dimension. The negative correlation regarding age implies that the older the user, the more excellent (i.e. *good, professional, beautiful, expensive,* and *strong*) the icons were perceived, and the younger the user, the more inferior (i.e. *bad, unprofessional, ugly, cheap* and *weak*) the icons were perceived. This finding supports prior literature where younger audiences were found to be more critical towards GUI aesthetics than older audiences (Oyibo *et al.*, 2018). This might be explained by the notion that younger people tend to focus more on hedonic pleasure than older people (Hsieh *et al.*, 2004; Wallendorf and Arnould, 1988). The positive correlation concerning time suggests that the more time the user spends interacting with the interface, the less appealing the icons were rated. With the increase of time, users will naturally adapt to icon aesthetics that essentially repeat similar patterns, which may lead to developing a critical eye towards graphical elements. This way the users establish a taste for iconography over time, which might make users more selective.

Interaction effects were found between age and gender, age and time, as well as age, gender and time for the VISQUAL models (Tables 5 and 7) on the dimensions Excellence/Inferiority and Normalness/Bizarreness. On the basis of the positive correlation between age and gender, especially younger male users perceived the icons as excellent, and especially older female users perceived the icons as inferior. This finding is similar to prior literature in the way that women have been shown to appreciate aesthetics more than men, and might thus be more critical towards design aspects (Creusen, 2010; Oyibo and Vassileva, 2017). Likewise, early technology adoption within younger age groups and especially men (Morris and Venkatesh, 2000; Venkatesh *et al.*, 2000) might lead to better ratings, as they are perhaps generally more used to viewing game app icons. The positive interaction between age and time on the Normalness/Bizarreness dimension suggests that the perception of normalness (i.e. *concrete*, *realistic*, *ordinary*) tends to increase with time spent interacting with interfaces. The negative three-way interaction between age, gender and time on the same dimension further indicates that especially younger women evaluated icons as more normal when more time was spent using app stores. This suggests that icon aesthetics might be difficult to grasp in the beginning, which eventually changes as the user continues interacting with the interface. As noted previously, users tend to adapt to icon aesthetics thus losing some of their perceived uniqueness.

When examining the results concerning age, gender and time with individual adjective pairs as dependent variables (Table 8), age affected *expensive–cheap*, *strong–weak*,

	Beta	*t*	*p*	Beta	*t*	*p*	Beta	*t*	*p*
		Age			Gender[a]			Time	
Excellence–Inferiority	*−0.170**	*−2.393*	*0.017*	0.002	0.069	0.945	*0.173**	*2.046*	*0.041*
Graciousness–Harshness	0.071	1.001	0.317	0.023	1.004	0.316	0.035	0.407	0.684
Idleness–Liveliness	−0.044	−0.616	0.538	−0.030	−1.309	0.191	0.107	1.256	0.209
Normalness–Bizarreness	0.037	0.520	0.603	−0.020	−0.908	0.364	0.013	0.155	0.877
Complexity–Simplicity	0.067	0.946	0.344	0.029	1.297	0.195	0.105	1.241	0.215

Note(s): * = $p < 0.05$, ** = $p < 0.01$, statistically significant effects italicized. [a]Females were coded with the higher variable value

Table 6.
Regression analyses (22 items) with age, gender and time spent browsing app stores (per week)

Table 7.
Regression analyses (22 items) with Age × Gender, Age × Time, Gender × Time, and Age × Gender × Time

	Beta	*t*	*p*	Beta	*t*	*p*	Beta	*t*	*p*	Beta	*t*	*p*
	Age × Gender			Age × Time			Gender × Time			Age × Gender × Time		
Excellence–Inferiority	*0.193***	*2.713*	*0.007*	0.026	0.329	0.742	−0.137	−1.602	0.109	0.006	0.073	0.942
Graciousness–Harshness	−0.053	−0.744	0.457	0.014	0.170	0.865	−0.038	−0.444	0.657	−0.002	−0.019	0.985
Idleness–Liveliness	0.068	0.949	0.343	0.102	1.276	0.202	−0.084	−1.041	0.150	−0.084	−1.041	0.298
Normalness–Bizarreness	−0.057	−0.798	0.425	*0.208***	*2.604*	*0.009*	−0.048	−0.556	0.578	*−0.182**	*−2.248*	*0.025*
Complexity–Simplicity	−0.065	−0.912	0.362	−0.074	−0.925	0.355	−0.060	−0.693	0.489	0.072	0.887	0.375

Note(s): * = $p \leq 0.05$, ** = $p < 0.01$, statistically significant effects italicized. [a]Females were coded with the higher variable value

Table 8. Regression analyses with age, gender and time spent browsing app stores (per week)

	Beta	*t* Age	*p*	Beta	*t* Gender[a]	*p*	Beta	*t* Time	*p*
Expensive–Cheap	*−0.163**	*−2.288*	*0.022*	0.002	0.090	0.928	0.119	1.404	0.161
Strong–Weak	*−0.172**	*−2.423*	*0.015*	−0.024	−1.046	0.296	0.158	1.863	0.063
Professional-Unprofessional	*−0.164**	*−2.300*	*0.022*	0.018	0.817	0.414	0.149	1.754	0.080
Soft–Hard	*0.157**	*2.209*	*0.027*	*0.058**	*2.585*	*0.010*	−0.084	−0.988	0.323
Warm–Cool	0.053	0.739	0.460	*0.053**	*2.345*	*0.019*	0.031	0.367	0.714
Ordinary–Unique	0.108	1.518	0.129	0.009	0.420	0.675	*−0.175**	*−2.067*	*0.039*
Simple–Complex	0.116	1.633	0.103	0.026	1.150	0.250	0.037	0.432	0.666
Young–Old	0.018	0.246	0.806	−0.041	−1.817	0.069	0.126	1.481	0.139
Beautiful–Ugly	−0.083	−1.169	0.242	0.019	0.847	0.397	0.147	1.738	0.082
Good–Bad	−0.120	−1.688	0.092	−0.013	−0.574	0.566	0.140	1.650	0.099
Realistic–Unrealistic	−0.037	−0.521	0.603	−0.019	−0.852	0.395	0.119	1.405	0.160
Concrete–Abstract	0.008	0.109	0.913	−0.030	−1.342	0.180	0.075	2.130	0.377
Masculine–Feminine	−0.113	−1.591	0.112	−0.014	−0.602	0.547	0.064	0.757	0.449
Delicate–Rugged	0.059	0.824	0.410	0.022	0.978	0.328	0.020	0.233	0.816
Relaxed–Stiff	0.079	1.108	0.268	0.022	0.953	0.341	0.006	0.075	0.940
Colorful–Colorless	−0.074	−1.034	0.301	−0.022	−0.987	0.324	0.059	0.700	0.484
Three–Two−dimensional	−0.031	−0.436	0.663	0.008	0.376	0.707	0.081	0.959	0.338
Active–Passive	−0.117	−1.647	0.100	−0.010	−0.439	0.661	0.062	0.734	0.463
Fast–Slow	−0.043	−0.610	0.542	−0.026	−1.136	0.256	0.079	0.927	0.354
Happy–Sad	0.127	1.781	0.075	−0.025	−1.107	0.268	0.036	0.430	0.667
Exciting–Calm	−0.042	−0.594	0.553	−0.029	−1.295	0.196	0.048	0.569	0.570
Quiet–Loud	0.072	1.016	0.310	0.025	1.114	0.266	−0.031	−0.367	0.713

Note(s): * = $p < 0.05$, ** = $p < 0.01$, statistically significant effects italicized. [a]Females were coded with the higher variable value

Table 9.
Regression analyses with Age × Gender, Age × Time, Gender × Time, and Age × Gender × Time

	Beta	*t*	*p*	Beta	*t*	*p*	Beta	*t*	*p*	Beta	*t*	*p*
	Age × Gender			Age × Time			Gender × Time			Age × Gender × Time		
Expensive–Cheap	*0.177**	*2.488*	*0.013*	−0.025	−0.316	0.752	−0.088	−1.020	0.308	0.053	0.655	0.512
Strong–Weak	*0.203***	*2.857*	*0.004*	0.024	0.299	0.765	−0.122	−1.311	0.190	0.002	0.024	0.981
Professional-Unprofessional	*0.181**	*2.545*	*0.011*	−0.027	−0.339	0.735	−0.100	−1.166	0.244	0.056	0.686	0.493
Ordinary–Unique	*−0.156**	*−2.197*	*0.028*	0.155	1.941	0.052	0.127	1.485	0.138	−0.123	−1.520	0.129
Good–Bad	*0.147**	*2.062*	*0.039*	0.036	0.453	0.651	−0.130	−1.511	0.131	−0.016	−0.198	0.843
Concrete–Abstract	0.019	0.264	0.792	*0.170**	*2.130*	*0.033*	−0.108	−1.252	0.211	*−0.166**	*−2.050*	*0.040*
Exciting–Calm	0.070	0.979	0.328	0.140	1.758	0.079	−0.080	−0.925	0.355	*−0.162**	*−1.994*	*0.046*
Colorful–Colorless	0.104	1.464	0.143	−0.094	−1.175	0.240	−0.031	−0.356	0.722	0.102	1.252	0.211
Beautiful–Ugly	0.089	1.245	0.213	0.102	1.277	0.202	−0.135	−1.571	0.116	−0.072	−0.892	0.272
Simple–Complex	−0.086	−1.212	0.226	−0.066	−0.825	0.409	−0.018	−0.204	0.838	0.060	0.744	0.457
Realistic–Unrealistic	0.009	0.124	0.901	0.104	1.310	0.190	−0.108	−1.262	0.207	−0.083	−1.029	0.304
Young–Old	0.012	0.172	0.863	−0.019	−0.240	0.810	−0.091	−1.054	0.292	0.039	0.481	0.631
Active–Passive	0.111	1.562	0.118	0.088	1.101	0.271	−0.093	−1.078	0.281	−0.071	−0.869	0.385
Fast–Slow	0.048	0.679	0.497	0.066	0.826	0.409	−0.101	−1.175	0.240	−0.066	−0.808	0.419
Warm–Cool	−0.059	−0.830	0.407	0.043	0.535	0.593	−0.060	−0.700	0.484	−0.035	−0.430	0.667
Happy–Sad	−0.055	−0.771	0.441	−0.007	−0.086	0.931	−0.017	−0.201	0.841	0.025	0.309	0.758
Soft–Hard	−0.168	−2.360	0.323	0.064	0.807	0.420	0.024	0.283	0.777	−0.055	-0.674	0.500
Three–Two-dimensional	0.007	0.096	0.923	−0.021	−0.264	0.792	−0.049	−0.568	0.570	0.024	0.296	0.767
Masculine–Feminine	0.126	1.763	0.078	0.133	1.668	0.095	−0.073	−0.851	0.395	−0.121	−1.487	0.137
Delicate–Rugged	−0.061	−0.855	0.392	−0.094	−1.173	0.241	−0.009	−0.106	0.916	0.065	0.797	0.425
Relaxed–Stiff	−0.107	−1.496	0.135	0.027	0.339	0.735	0.013	0.146	0.884	−0.011	−0.130	0.897
Quiet–Loud	−0.063	−0.879	0.379	0.003	0.035	0.972	0.033	0.384	0.701	0.024	0.296	0.767

Note(s): * = $p \leq 0.05$, ** = $p < 0.01$, statistically significant effects italicized. [a]Females were coded with the higher variable value

professional–unprofessional and *soft–hard*. The negative correlation for *expensive–cheap*, *strong–weak* and *professional–unprofessional,* as well as the positive correlation for *soft–hard,* indicates that the younger the user, the cheaper, weaker, unprofessional and harder the icons seemed. This strengthens the finding that young people are critical towards GUI aesthetics and perhaps more used to seeing app icons in general, leading to relatively poor evaluations. Gender differences were found for the adjective pairs *soft–hard* and *warm–cool,* indicating that female users perceived icons as harder and cooler compared to male users. These findings show that icon aesthetics are in certain ways perceived more harshly by women than men, perhaps relating to prior findings of women shown to be more sensitive to visual complexity (Creusen, 2010; Reinecke and Gajos, 2014; Smith, 1995). Time spent browsing app stores affected the adjective pair *ordinary–unique*. Hence, the longer the users spend time browsing app stores, the more ordinary the icons were perceived. This supports the previous indication of users developing a critical eye over time.

Several interaction effects with individual adjective pairs as dependent variables (Table 9) were found. Concerning age and gender, the adjective pairs *expensive–cheap*, *strong–weak*, *professional–unprofessional*, *ordinary–unique* and *good–bad* were statistically significant. These findings suggest that younger men perceived the icons as ordinary, while older women perceived the icons as cheap, weak, unprofessional and bad. This, again, may refer to the prior findings of women shown to appreciate aesthetics more than men as well as men being more comfortable with technology and perhaps generally more used to viewing game app icons. Furthermore, concerning age and time, it was found that for younger users, the more time they spend browsing app stores, the more concrete the icons seem. Finally, the negative three-way interaction between age, gender and time for the adjective pairs *concrete–abstract* and *exciting–calm* indicates that younger women evaluated icons as more concrete and exciting when more time was spent using app stores. These results strengthen prior findings of female users preferring expressive aspects (i.e. beauty and emotional value) (Creusen, 2010; Henry, 2002; Oyibo and Vassileva, 2017; Tuch *et al.*, 2010; Wallendorf and Arnould, 1988).

5.1 Theoretical contributions

The growing need for adaptive and appealing user interfaces requires more work in understanding how perceptions and demographic factors affect user interface design. This study adds to the topic of interaction research, where usability has dominated research partly at the expense of aesthetic considerations (Tractinsky *et al.*, 2000).

The results of this study contradict prior research on demographic factors and interface systems (Leong *et al.*, 2013) in that gender differences do exist, although they seem to be minimizing among the younger generation. The variety in perceptions between genders can partly be explained by strategies according to the adaptive decision-making theory, where users choose designs by filtering choices based on subjective impressions of aesthetics. As hypothesized previously, the decision-maker's strategies are dependent on the individual's history, which contributes to differences in perceptions.

The findings of this study are consistent with prior literature in that younger audiences are somewhat critical towards GUI aesthetics (Creusen, 2010; Hsieh *et al.*, 2004; Morris and Venkatesh, 2000; Oyibo and Vassileva, 2017; Oyibo *et al.*, 2018; Venkatesh *et al.*, 2000; Wallendorf and Arnould, 1988). However, interestingly, as the sample in this study focused on younger age groups and no significant effect was found between gender and time using interfaces, it seems that the unisex pattern identified by prior research (Morris *et al.*, 2005) is continued: gender differences are not as visible among younger users as they have been among older users. Moreover, this study adds to the prior findings that time affects user perceptions of interface aesthetics (Lee and Koubek, 2011; Hartmann *et al.*, 2008) in such a way that users become more selective over time.

In terms of the adaptive decision-making theory, which has been advanced only minimally in the context of aesthetics in interaction design, users have been found to apply a tradeoff strategy by weighting different attributes of designs to an extent by the users' background (Hartmann *et al.*, 2007a). Drawing from this theoretical framework, as both younger women and men nowadays may have received greater exposure to technology compared with the older generation, the background of users seems to have become more homogenous, thus leading to similar perceptions of visual aesthetics according to the theory. Thus, the decision-making strategies seem to have unified as well, which may imply a change to the psychological patterns of making decisions altogether.

Previous studies that have employed adaptive decision-making theory in similar contexts (Hartmann *et al.*, 2007a, b, 2008) posit that aesthetics should be matched to user profile. We contribute to the literature by offering deeper insight on how the cultural atmosphere seems to be changing user preferences and decision-making behavior to predict their intention towards the judgment of aesthetic appeal, thus aiding scholars to revisit theories on decision-making and aesthetic appeal. Particularly, as the adaptive decision-making theory has not been applied widely and recently in similar studies, a critical approach could be adapted in order to systematically build the theory for the development of new hypotheses. In conclusion, we have demonstrated that aesthetics is a component of design quality that is susceptible to the user's decision-making strategies. Implementing this theoretical framework shows evidence that user perception is a complex construct that requires the understanding of deeper behavioral meaning.

5.2 Practical implications

This study adds to the existing literature of designing graphical user interface elements relating to demographic effects. Aesthetic appeal is a complex matter, nonetheless, some practical implications can be made on the basis of the findings.

Design implication 1: The results suggest that younger users in general, as well as older women, tend to be more critical towards icon aesthetics. Thus, in order to visually appeal to the tastes of younger audiences and women, focusing on creating high quality designs (i.e. *high graphical fidelity*) is recommended, as the hedonic aspects need to be catered to across these demographic factors.

Design implication 2: Expectedly, time affects perceptions in that novice users perceive icons as more excellent than experienced users. Therefore, in order to visually appeal to more experienced users, designers may have to put in more effort and creativity.

Design implication 3: Overall, gender differences among younger users seem to be minimizing and therefore gender-neutral options could be considered in future design processes. However, the perceptions of icons change especially for younger women in that icons are seen as more concrete and exciting over time. Hence, practitioners could benefit from integrating young female users to interfaces at an early stage to increase the aforementioned effects.

The results suggest that user perceptions are subjective and thus age, gender and time have relatively little effect on how users evaluate icon aesthetics. However, these implications inform what kinds of aesthetic perceptions graphical user interface elements (i.e. *icons*) should be brought to evoke. This knowledge can then be adapted in establishing segmentation models for the design of adapted user interfaces.

5.3 Limitations and future research

This study was one of the first attempts to understand how demographic factors affect user perceptions of GUI element aesthetics by utilizing game app icons as data collection material. However, there are some limitations that should be acknowledged.

Game app icons were used as study material to maximize internal validity. This poses a possibility for conducting future research on other app icon types for comparative results. The choice of not informing participants about the purpose of the apps behind the icons was made to avoid systematic bias. However, it would be beneficial to conduct a similar study with additional information on the app context.

The data was gathered via an online survey that was advertised on Finnish student organizations' mailing lists, thus the sample can be considered fairly homogenous. The majority of the respondents are from the same age group and come from a similar cultural background, which could affect perceptions in the study. Moreover, the sample in this study is a nonprobability convenience sample, therefore it is not necessarily representative of all app store users. In future research, a more diverse sample should be gathered in order to gain perspective on factors related to age and cultural differences.

As is commonplace within the industry, actual data on app store usage was not available, thus the measurement used in this study reports intended behavior with a vignette style experiment setting. This may have an impact on the generalizability of the findings. Moreover, as the results consist of perceptions measured by quantitative means, the findings may be considered ambiguous with underlying biases. Therefore, a qualitative approach would be beneficial in order to gain a deeper understanding of the topic in further studies. Additionally, an even more authentic experiment setting could be composed.

6. Conclusion

This study replicated prior literature in the sense that paying close attention to visual aesthetics is important, especially when targeting experienced users, young audiences and women. Knowledge about demographic effects relating to how GUI elements (i.e. *app icons*) are perceived is scarce, therefore, insight into the topic is valuable for deciding on effective design processes. Considering the changing cultural atmosphere, especially relating to gender and age in the domain of technology, insight into the topic is valuable. The current undertaking shows that technology adoption advances at a tremendous pace, which blurs the boundaries of aesthetics between people despite their age, gender and habits in daily life.

References

Ahmed, S.U., Al Mahmud, A. and Bergaust, K. (2009), "Aesthetics in human-computer interaction: views and reviews", in Jacko, J.A. (Ed.), *Human-Computer Interaction, New Trends, HCI 2009, Lecture Notes in Computer Science*, Springer, Berlin, Vol. 5610.

Aiken, L.S. and West, S.G. (1991), *Multiple Regression: Testing and Interpreting Interactions*, Sage, London.

Arend, U., Muthig, K.P. and Wandmacher, J. (1987), "Evidence for global superiority in menu selection by icons", *Behaviour and Information Technology*, Vol. 6 No. 4, pp. 411-426, doi: 10.1080/01449298708901853.

Beresford, B. and Sloper, T. (2008), *Understanding the Dynamics of Decision-Making and Choice: A Scoping Study of Key Psychological Theories to Inform the Design and Analysis of the Panel Study*, Social Policy Research Unit, University of York, York.

Blankenberger, S. and Hahn, K. (1991), "Effects of icon design on human-computer interaction", *International Journal of Man-Machine Studies*, Vol. 35 No. 3, pp. 363-377, doi: 10.1016/S0020-7373(05)80133-6.

Boiano, S., Borda, A., Bowen, J., Faulkner, X., Gaia, G. and Mcdaid, S. (2006), "Gender issues in HCI design for web access", in Kurniawan, S. and Zaphiris, P. (Eds), *Advances in Universal Web Design and Evaluation: Research, Trends and Opportunities*, IGI Global, London, pp. 116-153.

Chartboost (2015), "Power-up report–July 2015", available at: https://chartboost.s3.amazonaws.com/blog/power-up-report-july-2015-building-an-empire-mobile-strategy-games.pdf (accessed 14 September 2018).

Choi, J.H. and Lee, H.J. (2012), "Facets of simplicity for the smartphone interface: a structural model", *International Journal of Human-Computer Studies*, Vol. 70 No. 2, pp. 129-142, doi: 10.1016/j.ijhcs.2011.09.002.

Chung, J.E., Park, N., Wang, H., Fulk, J. and McLaughlin, M. (2010), "Age differences in perceptions of online community participation among non-users: an extension of the technology acceptance model", *Computers in Human Behavior*, Vol. 26 No. 6, pp. 1674-1684.

Creusen, M.E.H. and Schoormans, J.P.L. (2005), "The different roles of product appearance in consumer choice", *Journal of Product Innovation Management*, Vol. 22 No. 1, pp. 63-81, doi: 10.1111/j.0737-6782.2005.00103.x.

Creusen, M.E.H., Veryzer, R.W. and Schoormans, J.P.L. (2010), "Product value importance and consumer preference for visual complexity and symmetry", *European Journal of Marketing*, Vol. 44 Nos 9/10, pp. 1437-1452, doi: 10.1108/03090561011062916.

Creusen, M.E.H. (2010), "The importance of product aspects in choice: the influence of demographic characteristics", *Journal of Consumer Marketing*, Vol. 27 No. 1, pp. 26-34, doi: 10.1108/07363761011012921.

Cyr, D. (2009), "Gender and website design across cultures", *Paper Presented at the 17th European Conference on Information Systems*, Verona, Italy, pp. 279-291.

Debevc, M., Meyer, B., Donlagic, D. and Svecko, R. (1996), "Design and evaluation of an adaptive icon toolbar", *User Modeling and User-Adapted Interaction*, Vol. 6 No. 1, pp. 1-21, doi: 10.1007/BF00126652.

Dewar, R. (1999), "Design and evaluation of public information symbols", in Zwaga, H.J.G., Boersema, T. and Hoonhout, H.C.M. (Eds), *Visual Information for Everyday Use*, Taylor & Francis, London, pp. 285-303.

Gait, J. (1985), "An aspect of aesthetics in human-computer communications: pretty windows", *IEEE Transactions on Software Engineering*, Vol. 11 No. 8, pp. 714-717, doi: 10.1109/TSE.1985.232520.

Genuine (2013), "Gender-inclusive user interface guidelines", available at: http://genuine.ict.tuwien.ac.at/unterlagen/Guidelines.pdf (accessed 27 May 2019).

Goonetilleke, R.S., Shih, H.M., On, H.K. and Fritsch, J. (2001), "Effects of training and representational characteristics in icon design", *International Journal of Human-Computer Studies*, Vol. 55 No. 5, pp. 741-760, doi: 10.1006/ijhc.2001.0501.

Hartmann, J., Sutcliffe, A. and De Angeli, A. (2007a), "Towards a theory of user judgment of aesthetics and user interface quality", *ACM Transactions on Computer-Human Interaction*, Vol. 15 No. 4, doi: 10.1145/1460355.1460357.

Hartmann, J., Sutcliffe, A. and De Angeli, A. (2007b), "Investigating attractiveness in web user interfaces", *Paper Presented at the 25th Annual SIGCHI Conference on Human Factors in Computing Systems*, San Jose, USA, pp. 387-396.

Hartmann, J., De Angeli, A. and Sutcliffe, A. (2008), "Framing the user experience: information biases on website quality judgement", *Paper Presented at the 26th Annual SIGCHI Conference on Human Factors in Computing Systems*, Florence, Italy, pp. 855-864.

Hassenzahl, M., Burmester, M. and Koller, F. (2003), "AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität [AttracDiff: a questionnaire to measure perceived hedonic and pragmatic quality]", in Ziegler, J. and Szwillus, G. (Eds), *Mensch and Computer 2003–Interaktion in Bewegung*, B. G. Teubner, Stuttgart, pp. 187-196.

Henry, P. (2002), "Systematic variation in purchase orientations across social classes", *Journal of Consumer Marketing*, Vol. 19 No. 5, pp. 424-438, doi: 10.1108/07363760210437641.

Hoffman, L.W. (1972), "Early childhood experiences and women's achievement motives", *Journal of Social Issues*, Vol. 28 No. 2, pp. 129-155.

Hou, K.-C. and Ho, C.-H. (2013), "A preliminary study on aesthetic of apps icon design", *Paper Presented at the 5th International Congress of International Association of Societies of Design Research*, Tokyo, Japan, pp. 3845-2856.

Hsiao, K.-L. and Chen, C.-C. (2016), "What drives in-app purchase intention for mobile games? An examination of perceived values and loyalty", *Electronic Commerce Research and Applications*, Vol. 16, pp. 18-29, doi: 10.1016/j.elerap.2016.01.001.

Hsieh, M.-H., Pan, S.-L. and Setiono, R. (2004), "Product-, corporate-, and country-image dimensions and purchase behavior: a multicountry analysis", *Journal of the Academy of Marketing Science*, Vol. 32 No. 3, pp. 251-270, doi: 10.1177/0092070304264262.

Huang, S.M., Shieh, K.K. and Chi, C.F. (2002), "Factors affecting the design of computer icons", *International Journal of Industrial Ergonomics*, Vol. 29 No. 4, pp. 211-218, doi: 10.1016/S0169-8141(01)00064-6.

Huang, S. (2013), "Usability and GUI design and principles", available at: https://www.uio.no/studier/emner/matnat/ifi/INF5120/v13/undervisningsmateriale/f04-2013_0402-3-designguidelines_additional_info.pdf (accessed 27 May 2019).

Isherwood, S.J., McDougall, S.J.P. and Curry, M.B. (2007), "Icon identification in context: the changing role of icon characteristics with user experience", *Human Factors*, Vol. 49 No. 3, pp. 465-476, doi: 10.1518/001872007X200102.

Jennings, M. (2000), "Theory and models for creating engaging and immersive ecommerce websites", *Paper Presented at the 2000 ACM SIGCPR Conference on Computer Personnel Research*, New York, USA, pp. 77-85, doi: 10.1145/333334.333358.

Johnson, J. and Finn, K. (2017), *Designing User Interfaces for an Aging Population: Towards Universal Design*, Morgan Kaufmann, Amsterdam.

Jylhä, H. and Hamari, J. (2019), "An icon that everyone wants to click: how perceived aesthetic qualities predict app icon successfulness", *International Journal of Human-Computer Studies*, Vol. 130, pp. 73-85, doi: 10.1016/j.ijhcs.2019.04.004.

Jylhä, H. and Hamari, J. (2020), "Development of measurement instrument for visual qualities of graphical user interface elements (VISQUAL): a test in the context of mobile game icons", *User Modeling and User-Adapted Interaction*, Vol. 30 No. 5, pp. 949-982, doi: 10.1007/s11257-020-09263-7.

KnowItAll Ninja (2016), "Exploring user interface design principles and project planning techniques", available at: https://www.pearsonschoolsandfecolleges.co.uk/AssetsLibrary/SECTORS/FurtherEducationColleges/SUBJECT/dit/dit-student-book/btec-techaward-dit-sb-9781292208374.pdf (accessed 27 May 2019).

Kurosu, M. and Kashimura, K. (1995), "Apparent usability vs. inherent usability", *Paper Presented at the CHI 95 Conference Companion on Human Factors in Computing Systems*, New York, USA, pp. 292-293, doi: 10.1145/223355.223680.

Lavie, T. and Tractinsky, N. (2004), "Assessing dimensions of perceived visual aesthetics of web sites", *International Journal of Human-Computer Studies*, Vol. 60 No. 3, pp. 269-298, doi: 10.1016/j.ijhcs.2003.09.002.

Lee, S. and Koubek, R.J. (2011), "The impact of cognitive style on user preference based on usability and aesthetics for computer-based systems", *International Journal of Human-Computer Studies*, Vol. 27 No. 11, pp. 1083-1114, doi: 10.1080/10447318.2011.555320.

Leong, L.-Y., Ooi, K.-B., Chong, A.Y.-L. and Lin, B. (2013), "Modeling the stimulators of the behavioural intention to use mobile entertainment: does gender really matter?", *Computers in Human Behavior*, Vol. 29 No. 5, pp. 2109-2121, doi: 10.1016/j.chb.2013.04.004.

Lin, C.L. and Yeh, J.T. (2010), "Marketing aesthetics on the web: personal attributes and visual communication effects", *Paper Presented at the 5th IEEE International Conference on Management of Innovation and Technology*, IEEE, pp. 1083-1088.

Linux Information Project (2004), "GUI definition", available at: http://www.linfo.org/gui.html (accessed 28 May 2019).

Liu, D. and Guo, X. (2017), "Exploring gender differences in acceptance of mobile computing devices among college students", *Information Systems and E-Business Management*, Vol. 15 No. 1, pp. 197-223, doi: 10.1007/s10257-016-0315-x.

Maity, R., Uttav, A., Gourav, V. and Bhattacharya, S. (2015), "A non-linear regression model to predict aesthetic ratings of on-screen images", *Paper Presented at the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction*, OZCHI 2015, Parkville, pp. 44-52, doi: 10.1145/2838739.2838743.

McDougall, S.J.P. and Reppa, I. (2008), "Why do I like it? The relationships between icon characteristics, user performance and aesthetic appeal", *Paper Presented at the Human Factors and Ergonomics Society 52nd Annual Meeting*, New York, USA, pp. 1257-1261, doi: 10.1177/154193120805201822.

McDougall, S.J.P. and Reppa, I. (2013), "Ease of icon processing can predict icon appeal", in *Proceedings of the 15th International Conference on Human-Computer Interaction*, Las Vegas, pp. 575-584, doi: 10.1007/978-3-642-39232-0_62.

McDougall, S.J.P., Curry, M.B. and de Bruijin, O. (1999), "Measuring symbol and icon characteristics: norms for concreteness, complexity, meaningfulness, familiarity, and semantic distance for 239 symbols", *Behavior Research Methods, Instruments, and Computers*, Vol. 31 No. 3, pp. 487-519, doi: 10.3758/BF03200730.

McDougall, S.J.P., de Bruijn, O. and Curry, M.B. (2000), "Exploring the effects of icon characteristics on user performance: the role of icon concreteness, complexity, and distinctiveness", *Journal of Experimental Psychology: Applied*, Vol. 6 No. 4, pp. 291-306, doi: 10.1037/1076-898X.6.4.291.

McDougall, S.J.P., Reppa, I., Kulik, J. and Taylor, A. (2016), "What makes icons appealing? The role of processing fluency in predicting icon appeal in different task contexts", *Applied Ergonomics*, Vol. 55, pp. 156-172, doi: 10.1016/j.apergo.2016.02.006.

Miniukovich, A. and De Angeli, A. (2014), "Visual impressions of mobile app interfaces", *Paper Presented at the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*, Helsinki, Finland, pp. 31-40, doi: 10.1145/2639189.2641219.

Minton, H.L. and Schneider, F.W. (1980), *Differential Psychology*, Waveland Press, Prospect Heights, IL.

Morris, M.G. and Venkatesh, V. (2000), "Age differences in technology adoption decisions: implications for a changing work force", *Personnel Psychology*, Vol. 53 No. 2, pp. 375-403.

Morris, M.G., Venkatesh, V. and Ackerman, P.L. (2005), "Gender and age differences in employee decisions about new technology: an extension to the theory of planned behavior", *IEEE Transactions on Engineering Management*, Vol. 52 No. 1, pp. 69-84.

Moyes, J. and Jordan, P.W. (1993), "Icon design and its effect on guessability, learnability, and experienced user performance", in Alty, J.D., Diaper, D. and Gust, S. (Eds), *People and Computers VIII*, Cambridge University Society, Cambridge, pp. 49-59.

Ngo, D.C.L., Samsudin, A. and Abdullah, R. (2000), "Aesthetic measures for assessing graphic screens", *Journal of Information Science and Engineering*, Vol. 16 No. 1, pp. 97-116.

Ngo, D.C.L., Teo, L.S. and Byrne, J.G. (2003), "Modelling interface aesthetics", *Information Sciences*, Vol. 152 No. 1, pp. 25-46, doi: 10.1016/S0020-0255(02)00404-8.

Ngo, D.C.L. (2001), "Measuring the aesthetic elements of screen designs", *Displays*, Vol. 22 No. 3, pp. 73-78, doi: 10.1016/S0141-9382(01)00053-1.

Norman, D.A. (2004), *Emotional Design: Why We Love (Or Hate) Everyday Things*, Basic Books, New York.

Overby, E. and Sabyasachi, M. (2014), "Physical and electronic wholesale markets: an empirical analysis of product sorting and market function", *Journal of Management Information Systems*, Vol. 31 No. 2, pp. 11-46, doi: 10.2753/MIS0742-1222310202.

Oyibo, K. and Vassileva, J. (2017), "The interplay of aesthetics, usability and credibility in mobile website design and the moderation effect of gender", *Journal on Interactive Systems*, Vol. 8 No. 2, pp. 4-19.

Oyibo, K., Ali, Y.S. and Vassileva, J. (2016), "An empirical analysis of the perception of mobile website interfaces and the influence of culture", in Orji, R., Reisinger, M., Busch, M., Dijkstra, A., Stibe, A. and Tscheligi, M. (Eds), *Proceedings of the Personalization in Persuasive Technology Workshop, Persuasive Technology 2016*, Salzburg, Austria, pp. 44-56.

Oyibo, K., Adaji, I. and Vassileva, J. (2018), "The effect of age and information design on the perception of visual aesthetic", *Paper Presented at the British Human Computer Interaction Workshop*, Belfast, UK. doi: 10.14236/ewic/HCI2018.208.

Pappas, I.O., Mikalef, P., Giannakos, M.N. and Kourouthanassis, P.E. (2019), "Explaining user experience in mobile gaming applications: an fsQCA approach", *Internet Research*, Vol. 29 No. 2, pp. 293-314, doi: 10.1108/IntR-12-2017-0479.

Payne, J., Bettman, J. and Johnson, E. (1993), *The Adaptive Decision Maker*, Cambridge University Press, Cambridge. doi: 10.1017/CBO9781139173933.

Reinecke, K. and Gajos, K.Z. (2014), "Quantifying visual preferences around the world", *Paper Presented at the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*, ACM, pp. 11-20.

Rogers, Y. and Oborne, D.J. (1987), "Pictorial communication of abstract verbs in relation to human-computer interaction", *British Journal of Psychology*, Vol. 78 No. 1, pp. 99-112, doi: 10.1111/j.2044-8295.1987.tb02229.x.

Rousseau, G.K., Lamson, N. and Rogers, W.A. (1998), "Designing warnings to compensate for age-related changes in perceptual and cognitive abilities", *Psychology and Marketing*, Vol. 15 No. 7, pp. 643-662.

Salimun, C., Purchase, H.C., Simmons, D. and Brewster, S. (2010), "The effect of aesthetically pleasing composition on visual search performance", *Paper Presented at the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, Reykjavik, Iceland, ACM, pp. 422-431, doi: 10.1145/1868914.1868963.

Salman, Y.B., Kim, Y. and Cheng, H. (2010), "Senior-friendly icon design for the mobile phone", *Paper Presented at the 6th International Conference on Digital Content, Multimedia Technology and its Applications (IDC 2010)*, IEEE, Seoul, pp. 103-108.

Sarsam, S.M. and Al-Samarraie, H. (2018), "Towards incorporating personality into the design of an interface: a method for facilitating users' interaction with the display", *User Modeling and User-Adapted Interaction*, Vol. 28 No. 1, pp. 75-96, doi: 10.1007/s11257-018-9201-1.

Shaikh, A.D. (2009), "Know your typefaces! Semantic differential presentation of 40 onscreen typefaces", *Usability News*, Vol. 11, pp. 23-65.

Shu, W. and Lin, C.-S. (2014), "Icon design and game app adoption", in *Proceedings of the 20th Americas Conference on Information Systems*, Georgia.

Smith, G.E. (1995), "Framing product design: using design communication to facilitate user learning", *Journal of Business and Industrial Marketing*, Vol. 10 No. 5, pp. 6-21, doi: 10.1108/08858629510103860.

Sun, H. and Zhang, P. (2006), "The role of moderating factors in user technology acceptance", *International Journal of Human-Computer Studies*, Vol. 64 No. 2, pp. 53-78.

Sweet, S.A. and Grace-Martin, K.A. (2011), *Data Analysis with SPSS: A First Course in Applied Statistics*, 4th ed., Pearson, London.

Thüring, M. and Mahlke, S. (2007), "Usability, aesthetics and emotions in human–technology interaction", *International Journal of Psychology*, Vol. 42 No. 4, pp. 253-264, doi: 10.1080/00207590701396674.

Tractinsky, N., Katz, A.S. and Ikar, D. (2000), "What is beautiful is useable", *Interacting with Computers*, Vol. 13 No. 2, pp. 127-145, doi: 10.1016/S0953-5438(00)00031-X.

Tuch, A.N., Bargas-Avila, J.A. and Opwis, K. (2010), "Symmetry and aesthetics in website design: it's a man's business", *Computers in Human Behavior*, Vol. 26 No. 6, pp. 1831-1837, doi: 10.1016/j.chb.2010.07.016.

Venkatesh, V. and Morris, M.G. (2000), "Why don't men ever stop to ask for directions? Gender, social influence, and their role in technology acceptance and usage behavior", *MIS Quarterly*, Vol. 24 No. 1, pp. 115-139.

Venkatesh, V., Morris, M.G. and Ackerman, P.L. (2000), "A longitudinal field investigation of gender differences in individual technology adoption decision-making processes", *Organizational Behavior and Human Decision Processes*, Vol. 83 No. 1, pp. 33-60.

Wallendorf, M. and Arnould, E.J. (1988), "'My favorite things': a cross-cultural inquiry into object attachment, possessiveness, and social linkage", *Journal of Consumer Research*, Vol. 14 No. 4, pp. 531-547.

Zen, M. and Vanderdonckt, J. (2016), "Assessing user interface aesthetics based on the inter-subjectivity of judgment", *Paper Presented at the 30th International BCS Human Computer Interaction Conference*, BCS, Swindon. doi: 10.14236/ewic/HCI2016.25.

Corresponding author
Henrietta Jylhä can be contacted at: henrietta.jylha@gmail.com