International Journal of Human-Computer Studies 130 (2019) 73–85

Contents lists available at ScienceDirect

International Journal of Human-Computer Studies

journal homepage: www.elsevier.com/locate/ijhcs

An icon that everyone wants to click: How perceived aesthetic qualities predict app icon successfulness☆

Henrietta Jylhä[a,*], Juho Hamari[a,b]

[a] Gamification Group, Faculty of Information Technology and Communication Sciences, Tampere University, 33014 Tampere University, Finland
[b] Gamification Group, Faculty of Humanities, University of Turku, Finland

ARTICLE INFO

Keywords:
Iconography
Aesthetics
Semantic differential
Mobile apps
Graphical user interfaces
Digital marketing

ABSTRACT

Mobile app markets have been touted as fastest growing marketplaces in the world. Every day thousands of apps are published to join millions of others on app stores. The competition for top grossing apps and market visibility is fierce. The way an app is visually represented can greatly contribute to the amount of attention an icon receives and to its consequent commercial performance. Therefore, the *icon* of the app is of crucial importance as it is the first point of contact with the potential user/customer amidst the flood of information. Those apps that fail to arouse attention through their icons danger their commercial performance in the market where consumers browse past hundreds of icons daily. Using semantic differential scale (22 adjective pairs), we investigate the relationship between consumer perceptions of app icons and icon successfulness, measured by 1) overall evaluation of the icon, 2) willingness to click the icon, 3) willingness to download the imagined app and, 4) willingness to purchase the app. The study design was a vignette study with random participant (n = 569) assignment to evaluate 4 icons (n = 2276) from a total of pre-selected 68 game app icons across 4 categories (concrete, abstract, character and text). Results show that consumers are more likely to interact with app icons that are aesthetically pleasing and convey good quality. Particularly, app icons that are perceived unique, realistic and stimulating lead to more clicks, downloads and purchases.

1. Introduction

After app stores became prominent venues for providing software, the number of mobile apps has been constantly growing at a fast pace (Moreira et al., 2014). Online storefronts try to attract critical masses in various ways, but effective design is necessary for consumer engagement (Overby and Sabyasachi, 2014). Rapid changes in the app markets and consumer mindsets poise new possibilities and challenges in the world-wide competition of commercial success, which motivates the need for further research on app icons and consumer behavior.

App stores house a massive number of mobile applications, also known as apps. To this date, the total number of app downloads from app stores worldwide is estimated 197 billion.[1] Furthermore, global apps industry revenue has been predicted to rise to 188.9 billion U.S. dollars in 2020.[2] In light of these statistics, the impact of the apps industry to economic growth is undeniably high. All apps are listed on app stores as icons – a graphic that "provides a quick, intuitive representation of an action, a status or an app".[3] An icon-based graphical user interface (GUI) common to smartphones and tablets has a limited display area, which is why app icons should provide good recognition and user preference (Böhmer and Krüger, 2013; Chen, 2015; Hou and Ho, 2013). Icons essentially act as a first-pass filter for saturated app markets, which is why they need to immediately capture a consumer's attention.[4] App icon is in many cases the first and most powerful opportunity to succeed in user engagement on the highly competitive app store markets (Woolridge and Schneider, 2011), hence developers and

☆ No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work.

* Corresponding author.
E-mail addresses: henrietta.jylha@tuni.fi (H. Jylhä), juho.hamari@tuni.fi (J. Hamari).
URL: http://gamification.group (H. Jylhä).

[1] Statista, "Number of mobile app downloads worldwide in 2016, 2017 and 2021 (in billions)," http://www.statista.com/statistics/234649/percentage-of-us-population-that-play-mobile-games/ (Accessed January 30, 2018).

[2] Statista, "Worldwide mobile app revenues in 2015, 2016 and 2020 (in billion U.S. dollars)," https://www.statista.com/statistics/269025/worldwide-mobile-app-revenue-forecast/ (Accessed January 30, 2018).

[3] Android Developers, "Iconography," http://www.androiddocs.com/design/style/iconography.html (Accessed January 30, 2018).

[4] Apple Developers, "App icon," https://developer.apple.com/ios/human-interface-guidelines/icons-and-images/app-icon/ (Accessed January 30, 2018).

https://doi.org/10.1016/j.ijhcs.2019.04.004
Received 3 May 2018; Received in revised form 23 April 2019; Accepted 24 April 2019
Available online 26 April 2019

designers must make a strong impact to prompt consumers to choose to interact with their app instead of the many others.

This observation leads us to the following key research questions: *How do consumer perceptions of app icon aesthetics affect icon successfulness*, namely, what are the aesthetic qualities that are likely to engage consumers into interacting with app icons? Moreover, does app icon appearance affect downloading and purchasing behavior of consumers? This topic is significant for research because minimal attention has been provided to how the visual attributes of apps represented on app stores affect consumer behavior (Wang and Li, 2017; Lin and Chen, 2018). Although icons appear commonly on various interfaces, research examining the determinants of icon appeal is scarce (McDougall et al., 1998, 2016). To our knowledge, no theoretical accounts have been proposed to explain the effects of consumer perceptions on app icon successfulness at the time of this research. Therefore, this study intends to lay the groundwork with potentially far-reaching practical and theoretical implications.

Using semantic differential scale (22 adjective pairs), this exploratory (i.e. non-confirmatory) study investigates the relationship between consumer perceptions of app icons and icon successfulness, measured by 1) overall evaluation of the icon, 2) willingness to click the icon, 3) willingness to download the imagined app and, 4) willingness to purchase the app. The study design is a vignette study, in which participants (n = 569) were assigned to evaluate 4 randomized icons from a total of pre-selected 68 game app icons across 4 categories (concrete, abstract, character and text). Game app icons were selected to maximize internal validity. This resulted in a total of 2276 individual icon evaluations. The findings show that consumers are more likely to interact with app icons that are aesthetically pleasing and convey good quality. This contrasts prior research (Shaikh, 2009) on onscreen typeface design and usage, which was implemented as a basis for our experiment. Furthermore, gaps in prior icon design theories are exposed regarding predictors of icon successfulness in terms the adjective pairs Concrete–Abstract and Complex–Simple, as no consistent statistically significant effect was found among them in our study.

2. Background

2.1. Graphical presentation in human–computer interaction

Icons are pictographic symbols of data or processes within a computer system, applied principally to graphics-based interfaces of operating systems (Gittins, 1986). Icons are widely used in human–computer interaction and they have replaced commands and menus as the means by which the computer supports a dialogue with the end-user (García et al., 1994; Gittins 1986; McDougall et al., 1998; Huang et al., 2002). Similar to mobile platforms, iconic interfaces have made their way into our everyday life. Advances in technology result in additional features and further, additional icons. The evolution of icons is traced back to signs (Goonetilleke et al., 2001). Signs are elements that "stand to someone for something in some respect or capacity" (Peirce, 1932). This can be interpreted in the sense that signs as well as icons have a symbolic meaning or connotation behind them. Prior research (Wiedenbeck, 1999) supports this by noting that icons are interface objects that represent a larger system in a simplified, pictorial manner. As we communicate through symbols, these symbols must also be embedded in icons to evoke the desired connotation in the viewer (Horton, 1996).

The terms icon and symbol are differentiated in that icons have a physical connection to a target or function, whereas symbols have an arbitrary, indirect relationship to that which they refer (Horton, 1994). However, the use of the term "icon" to describe symbols has become dominant especially in the interactive field (Horton, 1996). Thus, the everyday usage of "icon" stands for any graphic on an interactive button, and these icons can represent system objects such as files or folders, or actions such as messaging or calling (Wiedenbeck, 1999). Furthermore, leisurely icons, such as game and movie icons, often depict characters and other relevant features to the title which they represent. This is believed to enhance product identity and brand personality (Phillips, 1996).

The reason why icons are extensively used is due to many factors. Icons facilitate human–computer interaction because they are swiftly recognized and memorized (Horton, 1994, 1996; McDougall et al., 1999; Wiedenbeck, 1999). Icons are also more convenient for universal communication than text, since language interpretation is not an obstacle (Arend et al., 1987; Horton, 1994, 1996; Lodding, 1983; McDougall et al., 1999). Despite these positive results of icon usage, there is little published research on app icons, justifying further investigation.

One aspect of prior research on icon aesthetics concerns whether concrete or abstract icons are more effective from user perspective (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., 1998, 1999, 2000; Moyes and Jordan, 1993; Rogers and Oborne, 1987). Icon concreteness is the extent to which it depicts real objects (Isherwood et al., 2007), whereas icon abstractness tends to have less obvious connections with real objects (McDougall et al., 1999). Some studies (e.g. Hou and Ho, 2013; McDougall and Reppa, 2008; McDougall et al., 1999, 2000; Moyes and Jordan, 1993; Rogers and Oborne, 1987) show that most users prefer concrete, rich icon designs to abstract, simplified icons, while others have found conflicting results (Arend et al., 1987; McDougall et al., 1998). Prior research has also suggested that concreteness may not be of primary importance after all, rather semantic distance and familiarity may be more important (Blankenberger and Hahn, 1991; Dewar, 1999; Isherwood et al., 2007; McDougall et al., 1998, 1999; McDougall and Reppa, 2008; Schröder and Ziefle, 2008). Furthermore, icon familiarity has been acknowledged to help reduce the amount of information to communicate a message (Arab et al., 2013; Forsythe et al., 2008) which makes an icon easier to understand.

The juxtaposition of concrete and abstract icons is sometimes referred to as the *guessability gulf* (Moyes and Jordan, 1993). This is because concrete icons are easier to cognitively process at first sight than abstract icons. Despite the debate between concreteness and abstractness of icons, it is noteworthy that icon preference is affected by many factors. Computer icons have evolved from information signs to a part of consumer culture (Huag et al., 2002), therefore different types of icons may be suitable for different purposes and personalities. For example, concrete icons can be useful in public information systems or warnings (McDougall et al., 1998; McDougall and Reppa, 2013) where the goal is to clearly communicate information, whereas more stylistic icon design may promote other ends (Hou and Ho, 2013).

Another aspect of effective icon design is the speed and ease with which icons can be understood (e.g. Arend et al., 1987; Blankenberger and Hahn, 1991; Dewar, 1999; Lodding, 1983; Isherwood et al., 2007; McDougall and Reppa, 2008; McDougall and Reppa, 2013; McDougall et al., 2016; Moyes and Jordan, 1993; Wiedenbeck, 1999). Prior research (McDougall and Reppa, 2013, McDougall et al., 2016) on interface icon design has found that processing fluency affects icon appeal and that simple icon design has been shown to lead to user satisfaction. Factors influencing icon processing are e.g. icon familiarity and complexity, meaning that the easier the icon is to process due to simple design and earlier experience with similar icons, the more appealing it is (Arab et al., 2013; Choi and Lee, 2012; Dewar, 1999; Forsythe et al., 2008; Goonetilleke et al., 2001; Huang et al., 2002; Isherwood et al., 2007; McDougall and Reppa, 2008; McDougall et al., 1998; 2016; McDougall and Reppa, 2013; Moyes and Jordan, 1993).

Further concerning icon design and icon effectiveness, it has been speculated by prior research that the most important features of an icon are dominance, uniqueness or discriminability, and unambiguity (Arend et al., 1987; Dewar, 1999; Goonetilleke et al., 2001; Huang et al., 2002). Dominance is defined as a characteristic intrinsic to a

function and its context. An icon element is dominant if the other icon elements can be inferred from the first one. Therefore, icons with dominant elements are said to be processed more fluently than icons with redundant elements (Goonetilleke et al., 2001). An icon is said to convey uniqueness or discriminability when the representation and its function has perceptual immediacy, making an icon distinguishable and locatable among other icons (Goonetilleke et al., 2001; Huang et al., 2002). For example, icons featuring elements that are not interchangeable with other representations enhance icon uniqueness. The use of unique, visually distinctive icons has been shown to lead into better performance compared to icons that are not perceived as unique (Arend et al., 1987; Goonetilleke et al., 2001; Huang et al., 2002). However, uniqueness/discriminability has not been defined formally as specific differences in icons but rather it is in the eye of the beholder. Therefore, there is not a tangible way to define what it means in terms of specified features of an icon. Unambiguity is defined as a representation that can be associated with only one of the functions in a given context (Goonetilleke et al., 2001). If an icon is ambiguous, i.e. it holds multiple meanings in a single context, it may result in various interaction problems to users, especially if they have limited experience in icon identification (Black, 2017; Goonetilleke et al., 2001; Rogers and Oborne, 1987; Salman et al., 2010). For example, an icon depicting only a human face with an open mouth can refer to various actions such as eating, drinking or speaking, and therefore the representation is ambiguous. In this case, adding another element guiding the contextual function would aid in reaching unambiguity. Such an element could be e.g. a cup of coffee, which would make the icon easier to interpret as something related to the function of drinking (Goonetilleke et al., 2001).

App icons are a necessary part of branding and product design, as icons are key marketing elements presented to the consumer before downloading an app (i.e. product). Effective package and product design has been widely acknowledged as a factor for advantage in economic competition (e.g. Ares et al., 2011; Creusen and Schoormans, 2005; Creusen et al., 2010; Rundh, 2009; Schifferstein et al., 2013). Consumers use a lot of time and effort to evaluate how a product is presented, and often form their perceptions on brands based on design (Orth and Malkewitz, 2008). Hence, design affects brand and product selection and may drive purchase decisions (Ares et al., 2011; Creusen and Schoormans, 2005; Creusen et al., 2010; Fenko et al., 2010; Orth and Malkewitz, 2009; Schifferstein et al., 2013; van Rompay et al., 2009). In this light, the effects of product design should be of great importance to app designers, marketers and developers.

Product presentation can be divided into two main categories, visual and informational elements. Visual elements (i.e. graphics) include layout, color, typography, size and shape, whereas informational elements include written information about the product (Silayoi and Speece, 2004). App icons belong to the category of visual elements that communicate to the consumer most directly. In decision-making, consumers spontaneously form impressions of product content quality based on how a product is presented (Underwood et al., 2001; Yun et al., 2003) and these impressions can have lasting impact. For example, consumers perceive highly saturated colors as exciting (Labrecque and Milne, 2011), making them popular in product presentation. Furthermore, it is believed that effective visual elements in product presentation evoke more of an emotional response than informational elements (Silayoi and Speece, 2004), which in turn brings extra value to the product and increases the possibility of purchase (Cho and Lee, 2005). Hence, emotional impact is important when creating products, services and brands (Crossley, 2003).

2.2. Related work

Prior studies have investigated effective icon design in terms of icon concreteness and abstractness (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., 1999,2000; Moyes and Jordan, 1993; Rogers and Oborne, 1987), icon familiarity (Arab et al., 2013; Dewar, 1999; Forsythe et al., 2008; Huang et al., 2002; Isherwood et al., 2007; McDougall and Reppa 2008, McDougall and Reppa, 2013; McDougall et al., 2016; Moyes and Jordan, 1993), icon simplicity and complexity (Choi and Lee, 2012; Goonetilleke et al., 2001; McDougall and Reppa, 2008, McDougall and Reppa, 2013; McDougall et al., 2016) as well as uniqueness/discriminability and unambiguity (Creusen and Schoormans, 2005; Creusen et al., 2010; Dewar, 1999; Goonetilleke et al., 2001; Huang et al., 2002; Salman et al., 2010). However, app icons on app stores differ from other interface icons in that they are not only designed for interaction, but also as marketing assets meant to attract the consumer and to stand out in a display of many other offerings, much like any other, more tangible product. From this perspective, there are a plethora of other aspects that could also prove to be important determinant of app icon successfulness.

The handful of studies that have investigated the relationship between app icons and consumer behavior have consensus in that app icons play an important role in the mobile app markets, and that attractive icons have the power to trigger the interest of consumers (Burgers et al., 2016; Chen, 2015; Hou and Ho, 2013; Lin and Chen, 2018; Salman et al., 2010; Shu and Lin, 2014; Wang and Li, 2017). Nevertheless, mixed results have been reported on the attributes of successful icons. Studies that have investigated app icon design to understand task performance and user preference of different icon types have found that consumers prefer detailed, pictorial app icon design (Chen, 2015; Hou and Ho, 2013), sometimes regardless of inefficiency on task performance (Chen, 2015). Other findings on the visual attributes of app icon appearance recommend simplicity and complexity to be balanced for consumer appeal, as well as adding slight asymmetry to the design (Wang and Li, 2017). Moreover, positive evidence suggests that color is an important aspect of app icon design, as particularly bright and colorful icons increase the chance of app downloads and consumer preference (Salman et al., 2010; Wang and Li, 2017). Prior research on the relationship between icon attributes and consumer choice of apps has reported that app icon successfulness is dependent on app type as well as user personality and demographics (Hou and Ho, 2013; Salman et al., 2010; Shu and Lin, 2014), which complicates conclusions on the topic. Conflicting findings may be due to the fact that aesthetic appeal is a multi-dimensional topic that consists of various dimensions (Reppa and McDougall, 2015). According to the literature review herein, it appears that currently there does not exist a coherent body of knowledge on the issue of understanding how icon aesthetics affect perception and behavior. This is especially so as there exist only few studies on the topic as well as because their results are slightly mixed and conflicting. No clear trajectory of results emerges from the literature. Apps are used for several purposes by users with different profiles, thus it is important to advance knowledge in this topic to avoid pitfalls during icon design. As literature on this topic is limited, further investigation is justified.

Therefore, we set out to explore the relationship between consumer perceptions of app icon aesthetics and icon successfulness, to find out what are the perceived aesthetic qualities that are likely to engage consumers into interacting with app icons to fill the gap in prior literature identified above. Furthermore, we wish to further the body of research so that a pathway to a more coherent conclusion of icon aesthetics and consumer perceptions could be formed. The following section introduces the study design.

3. Methods and data

In order to find out how consumer perceptions of app icon aesthetics affect icon successfulness, we employed a semantic differential scale of 22 adjective pairs, measured by 1) overall evaluation of the icon, 2) willingness to click the icon, 3) willingness to download the imagined

Table 1
Demographic information.

		n	%
Age	–20	60	10.54
(SD = 7.24)	21–25	249	43.76
(Mean = 26.90)	26–30	145	25.48
(Median = 25.00)	31–35	45	7.91
	36–40	37	6.50
	41–45	16	2.81
	46–50	7	1.23
	51–55	5	0.88
	56–60	3	0.53
	60–	2	0.35
Education	Less than high school	5	0.9
	High school	135	23.7
	College	95	16.7
	Bachelor's degree	227	39.9
	Master's degree	98	17.2
	Higher than master's degree	9	1.6
Employment	Working full-time	133	23.4
	Working part-time	62	10.9
	Student	351	61.7
	Unemployed	11	1.9
	Retired	1	0.2
Game apps downloaded (per week)	0	429	75.4
	1	104	18.3
	2	14	2.5
	3	9	1.6
	4	2	0.4
	5	1	0.2
	Missing	10	1.8
Gender	Male	297	52.2
	Female	257	45.2
	Other	15	2.6
Yearly income	Less than $19,999	330	58.0
	$20,000 to $39,999	105	18.5
	$40,000 to $59,999	57	10.0
	$60,000 to $79,999	25	4.4
	$80,000 to $99,999	13	2.3
	$100,000 to $119,999	14	2.5
	$120,000 to $139,999	10	1.8
	$140,000 or more	15	2.6

app and, 4) willingness to purchase the app. This study utilized a within-subjects vignette approach, where each subject (n = 569) served in four treatments. Participants were assigned to evaluate 4 randomized icons from a total of pre-selected 68 game app icons across 4 categories (concrete, abstract, character and text) in a hypothetical situation setting instead of a description more typical to vignette studies. The aim was to acquire reliable data by exposing the participants close to a realistic setting outside the app store context. This resulted in a total of 2276 individual icon evaluations. The experiment was carried out as a self-administered online task. The following Section 3.1 describes the participants in the study.

3.1. Participants

The sample is composed of a nonprobability convenience sample with 569 respondents who participated in the study and assessed game app icons through the vignette experiment. A link to the online experiment was advertised in Facebook groups and Finnish student organizations' mailing lists. The participants predominantly resided in Finland (92.8%). Other countries clearly represented in the data were the United States (2.1%) and United Kingdom (2.1%). Please refer to Table 1 for demographic details of participants.

The gender split across participants was rather equal, as only slightly more than half were male (52.2%). The mean age was 26.90 years (SD = 7.24 years; 16–62 years). Most participants were university students (61.7%) and had a university-level education (39.9%). On a weekly basis, most participants (75.4%) did not download any game apps. Missing data (1.8%) was encountered for this item, as the frequency of mobile game downloads was only asked of 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 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, 2000; Moyes and Jordan, 1993; Rogers and Oborne, 1987), hence they were included in this experiment. The icons are presented in Table 2.

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.[5] 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. Since the categories are overlapping to an extent, separation between the categories was based on the most prominent elements in the icons. For example, icons in the "concrete" genre were selected in such a way that facial structures were not dominant in the icon, whereas in the "character" genre, the main element in the selected icons was a close-up image where the facial expression was prevalent.

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 3rd to 17th 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 sample of icons as possible, meaning that several different computer graphic techniques were included, such as 2D and 3D rendered images.

3.3. Measurements

Semantic differential scale was used to measure respondent

[5] Chartboost, "Power-Up Report – July 2015," https://chartboost.s3.amazonaws.com/blog/power-up-report-july-2015-building-an-empire-mobile-strategy-games.pdf (Accessed September 14, 2018).

Table 2
Icons in the study.

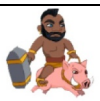
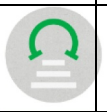
Category	Concrete	Abstract	Character	Text
Action				
Adventure				
Arcade				
Board				
Card				
Casino				
Casual				
Educational				
Music				
Puzzle				
Racing				
Role Playing				
Simulation				
Sports				
Strategy				
Trivia				
Word				

Table 3
Adjectives, means and standard deviations.

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

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, 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.

Table 3 lists the adjective pairs used in the study and presents an overview of the means and standard deviations. There were no outlier values and the range between the lowest and highest scores cluster closer 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.

To measure participants willingness to interact with the icons presented to them, 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

Table 4
The relationship between consumer perceptions of icons and the willingness to click, download and purchase.

	Beta Evaluation (R^2 = 0.658)	*P*	Beta Click (R^2 = 0.550)	*p*	Beta Download (R^2 = 0.530)	*p*	Beta Purchase (R^2 = 0.425)	*p*
Beautiful–Ugly	**−0.246****	**0.000**	**−0.256****	**0.000**	**−0.222****	**0.000**	**−0.201****	**0.000**
Good–Bad	**−0.332****	**0.000**	**−0.357****	**0.000**	**−0.351****	**0.000**	**−0.303****	**0.000**
Unique–Ordinary	**−0.071****	**0.000**	**−0.112****	**0.000**	**−0.098****	**0.000**	**−0.113****	**0.000**
Hard–Soft	**0.049****	**0.004**	**0.055****	**0.008**	**0.056****	**0.009**	**0.054***	**0.020**
Calm–Exciting	**0.072****	**0.000**	**0.069****	**0.002**	**0.086****	**0.000**	**0.049***	**0.043**
Passive–Active	**0.057****	**0.004**	**0.084****	**0.000**	**0.049***	**0.048**	0.029	0.276
Realistic–Unrealistic	−0.002	0.888	**−0.048****	**0.007**	**−0.052****	**0.004**	**−0.060****	**0.002**
Quiet–Loud	−0.013	0.462	**−0.057****	**0.007**	**−0.053***	**0.016**	**−0.051***	**0.033**
Colorful–Colorless	**−0.036***	**0.032**	**0.051***	**0.014**	0.030	0.156	**0.053***	**0.021**
Feminine–Masculine	**0.081****	**0.000**	**0.044***	**0.027**	0.037	0.068	0.021	0.328
Two–Three-dimensional	**0.031***	**0.036**	**−0.050****	**0.006**	−0.029	0.113	−0.007	0.719
Old–Young	**0.043****	**0.004**	0.020	0.256	0.027	0.147	0.014	0.485
Professional–Unprofessional	**−0.126****	**0.000**	−0.029	0.219	−0.048	0.051	−0.048	0.069
Relaxed–Stiff	**−0.055****	**0.002**	−0.013	0.554	−0.033	0.137	−0.035	0.148
Strong–Weak	**−0.060****	**0.000**	−0.027	0.194	−0.012	0.564	−0.020	0.396
Happy–Sad	0.002	0.907	0.023	0.275	0.042	0.053	**0.059***	**0.012**
Concrete–Abstract	0.024	0.118	0.015	0.413	0.031	0.103	0.039	0.057
Complex–Simple	0.004	0.800	−0.007	0.688	0.008	0.664	0.001	0.954
Cool–Warm	0.000	0.985	0.010	0.569	−0.002	0.911	0.013	0.489
Delicate–Rugged	−0.003	0.832	0.008	0.672	0.011	0.595	−0.001	0.980
Expensive–Cheap	−0.032	0.120	−0.005	0.829	−0.033	0.188	−0.025	0.354
Slow–Fast	−0.018	0.354	0.015	0.547	0.015	0.547	0.043	0.110

*$p < 0.05$, **$p < 0.01$, statistically significant effects bolded.

icon successfulness.

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 game 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 min 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

Multiple linear regression analyses (MLRA) were performed to investigate the relationships between perceptions of icons (represented by the 22 adjective pairs) and each of the four variables related to icon successfulness (1. overall evaluation of the icon, 2. willingness to click the icon, 3. willingness to download the imagined app and, 4. willingness to purchase the imagined app). Please refer to Table 4 for results.

We tested for multicollinearity with variance inflation factors (VIF), a common procedure in regression analysis to observe whether some relationships are masked due to collinearity. Multicollinearity causes inflation in the variances of regression coefficients, which may lead in unreliable conclusions about the relationship between variables. The VIF values for each adjective pair were between 1.5 and 2.7, except for the adjective pair Beautiful–Ugly and Good–Bad (VIF >3). Please refer to Table 5 for VIF scores.

According to Montgomery and Peck (1992) a VIF value that exceeds 5 (or in some cases 10) implies multicollinearity. In this light, the values in the analysis are more than acceptable. Nevertheless, compared to the other values, the higher VIF of Beautiful–Ugly and Good–Bad may suggest some multicollinearity, albeit that the values are not critically high. However, omitting variables due to relatively high VIF values (in comparison with other variables in the models) is a standard procedure that can be performed as a theory-driven decision. In this study, we aim to make predictions on the more underlying elements of icon aesthetics than those that are conceptually overlapping (i.e. Beautiful–Ugly and Good–Bad) at a higher level. Hence, we considered it worth finding out if there are significant elements hidden in the model when the adjective pairs of the highest VIF scores are removed. Thus, we ran additional post-hoc regression analyses excluding adjective pairs Beautiful–Ugly and Good–Bad. The analyses were performed with the remaining 20 adjective pairs and each of the four variables related to icon successfulness (1. overall evaluation of the icon, 2. willingness to click the icon, 3. willingness to download the imagined app and, 4. willingness to purchase the imagined app). Please refer to Table 6 for results.

Our predictions regarding hidden relationships between variables

Table 5
VIF values.

	VIF
Beautiful–Ugly	3.206
Good–Bad	3.494
Unique–Ordinary	1.326
Hard–Soft	1.924
Calm–Exciting	2.085
Passive–Active	2.570
Realistic–Unrealistic	1.368
Quiet–Loud	2.033
Colorful–Colorless	1.899
Feminine–Masculine	1.730
Two–Three-dimensional	1.443
Old–Young	1.420
Professional–Unprofessional	2.549
Relaxed–Stiff	2.065
Strong–Weak	1.922
Happy–Sad	1.963
Concrete–Abstract	1.503
Complex–Simple	1.338
Cool–Warm	1.350
Delicate–Rugged	1.760
Expensive–Cheap	2.725
Slow–Fast	2.579

due to multicollinearity were supported by the results of the post-hoc analyses (Table 6). Hidden significant effects were found when the analyses were performed without the adjective pairs Beautiful–Ugly and Good–Bad. This is probably caused by the general nature of the adjective pairs that may cause some of the relevant effects to remain undetected when they are kept in the model. Thus, in future research, the initial model could be corrected in such a way that the adjective pairs Beautiful–Ugly and Good–Bad are removed, as they may bias relationships with other variables.

The results indicate that regarding the relationship between consumer perceptions of app icons and their *overall evaluation*, the following ends of the semantic differentials positively predicted their grade (Table 4): *beautiful, good, unique, soft, exciting, active, colorful, masculine, three-dimensional, young, professional, relaxed* and *strong*. Without the adjective pairs Beautiful–Ugly and Good–Bad, the following adjectives also revealed to be positive predictors of the overall evaluation (Table 6): *expensive, quiet, realistic, happy, and simple*.

The following ends of the semantic differentials positively predicted *the willingness to click app icons* (Table 4): *beautiful, good, unique, soft, exciting, active, realistic, quiet, colorless, masculine* and *two-dimensional*. Without the adjective pairs Beautiful–Ugly and Good–Bad, the following adjectives also revealed to be positive predictors of the willingness to click (Table 6): *professional, expensive, strong, relaxed, happy, and young*.

The following ends of the semantic differentials positively predicted *the willingness to download the imagined app that the icon belongs to* (Table 4): *beautiful, good, unique, soft, exciting, active, realistic* and *quiet*. Without the adjective pairs Beautiful–Ugly and Good–Bad, the following adjectives also revealed to be positive predictors of the willingness to download (Table 6): *professional, expensive, strong, happy, young*, and *simple*.

The following ends of the semantic differentials positively predicted *the willingness to purchase the imagined app that the icon belongs to* (Table 4): *beautiful, good, unique, soft, exciting, realistic, quiet, colorless* and *sad*. Without the adjective pairs Beautiful–Ugly and Good–Bad, the following adjectives also revealed to be positive predictors of the willingness to purchase (Table 6): *professional, expensive, strong, relaxed,* and *fast*.

Additionally, both of the previous multiple linear regression models (Tables 4 and 6) were performed with stepwise method to acquire a more thorough understanding of the perceptions of icons (represented by the 22 adjective pairs) and each of the four variables related to icon successfulness (1. overall evaluation of the icon, 2. willingness to click the icon, 3. willingness to download the imagined app and, 4. willingness to purchase the imagined app). The purpose of these analyses was to observe whether there are differences in which adjective pairs are inserted in the models, in order to compare to the previous MLRA models. Please refer to Tables 7 and 8 for results.

The stepwise regressions inserted 12 to 9 of the original 22 adjective pairs in the final models per dependent (overall evaluation: 12, willingness to click: 11, willingness to download: 12, and willingness to purchase: 9). Nearly all variables in the models (except the adjective pair Happy–Sad within the willingness to download in Table 8) were statistically significant at the 0.05 or 0.01 level. These findings support the results of the previous regression analyses (Table 4) in that the adjective pairs inserted in the final models of the stepwise analyses were nearly identical to the previous MLRA.

In order to compare the results to the post-hoc MLRA (Table 6), we ran the stepwise regressions without the adjective pairs Beautiful–Ugly and Good–Bad to find out if hidden relationships remained. Please refer

Table 6
The relationship between consumer perceptions of icons and the willingness to click, download and purchase (excluding Beautiful–Ugly and Good–Bad).

	Beta	*p*	Beta	*p*	Beta	*p*	Beta	*p*
	Grade (R^2 = 0.567)		Click (R^2 = 0.521)		Download (R^2 = 0.506)		Purchase (R^2 = 0.408)	
Unique–Ordinary	**−0.101****	**0.000**	**−0.143****	**0.000**	**−0.128****	**0.000**	**−0.139****	**0.000**
Professional–Unprofessional	**−0.290****	**0.000**	**−0.204****	**0.000**	**−0.212****	**0.000**	**−0.192****	**0.000**
Expensive–Cheap	**−0.182****	**0.000**	**−0.165****	**0.000**	**−0.182****	**0.000**	**−0.157****	**0.000**
Hard–Soft	**0.065****	**0.001**	**0.071****	**0.002**	**0.071****	**0.002**	**0.067****	**0.006**
Strong–Weak	**−0.148****	**0.000**	**−0.120****	**0.000**	**−0.100****	**0.000**	**−0.097****	**0.000**
Relaxed–Stiff	**−0.111****	**0.000**	**−0.072****	**0.002**	**−0.088****	**0.000**	**−0.083****	**0.001**
Quiet–Loud	**−0.084****	**0.000**	**−0.133****	**0.000**	**−0.123****	**0.000**	**−0.112****	**0.000**
Calm–Exciting	**0.100****	**0.000**	**0.099****	**0.000**	**0.114****	**0.000**	**0.073****	**0.004**
Realistic–Unrealistic	**−0.050****	**0.002**	**−0.099****	**0.000**	**−0.100****	**0.000**	**−0.102****	**0.000**
Passive–Active	**0.078****	**0.000**	**0.106****	**0.000**	**0.070****	**0.009**	0.048	0.090
Happy–Sad	**−0.101****	**0.000**	**−0.086****	**0.000**	**−0.062****	**0.006**	−0.032	0.185
Old–Young	**0.060****	**0.000**	**0.039***	**0.048**	**0.043***	**0.028**	0.029	0.167
Colorful–Colorless	**−0.042***	**0.027**	**0.045***	**0.047**	0.025	0.281	**0.048***	**0.046**
Two–Three-dimensional	**0.038***	**0.022**	**−0.042***	**0.032**	−0.023	0.255	−0.001	0.951
Complex–Simple	**0.039***	**0.014**	0.031	0.106	**0.043***	**0.025**	0.032	0.113
Feminine–Masculine	**0.065****	**0.000**	0.027	0.214	0.023	0.295	0.008	0.715
Slow–Fast	0.000	0.982	0.034	0.192	0.033	0.212	**0.059***	**0.036**
Concrete–Abstract	0.029	0.084	0.021	0.293	0.036	0.073	0.044	0.042
Cool–Warm	0.006	0.721	0.016	0.395	0.004	0.847	0.018	0.364
Delicate–Rugged	−0.016	0.384	−0.005	0.831	−0.001	0.969	−0.011	0.635

*$p < 0.05$, **$p < 0.01$, statistically significant effects bolded.

Table 7
Overall evaluation and the willingness to click (stepwise).

Step #	Beta Evaluation (R^2 = 0.656)	p	Step #	Beta Click (R^2 = 0.490)	p
1 Good–Bad	−0.337**	0.000	1 Good–Bad	−0.376**	0.000
2 Beautiful–Ugly	−0.253**	0.000	2 Beautiful–Ugly	−0.273**	0.000
3 Passive–Active	0.049**	.004	3 Unique–Ordinary	−0.122**	0.000
4 Professional–Unprofessional	−0.142**	0.000	4 Passive–Active	0.095**	0.000
5 Unique–Ordinary	−0.079**	0.000	5 Colorful–Colorless	0.050**	0.008
6 Calm–Exciting	0.066**	0.000	6 Calm–Exciting	0.074**	0.000
7 Old–Young	0.047**	.001	7 Quiet–Loud	−0.053**	0.009
8 Feminine–Masculine	0.071**	0.000	8 Feminine–Masculine	0.052**	0.004
9 Relaxed–Stiff	−0.061**	0.000	9 Two–Three-dimensional	−0.047**	0.006
10 Strong–Weak	−0.060**	0.000	10 Realistic–Unrealistic	−0.040*	0.012
11 Hard–Soft	0.052**	.001	11 Hard–Soft	0.043*	0.015
12 Two–Three-dimensional	0.028*	.036			

* $p < 0.05$.
** $p < 0.01$.

Table 8
The willingness to download and purchase (stepwise).

Step #	Beta Download (R^2 = 0.466)	p	Step #	Beta Purchase (R^2 = 0.372)	p
1 Good–Bad	−0.363**	0.000	1 Good–Bad	−0.324**	0.000
2 Beautiful–Ugly	−0.233**	0.000	2 Beautiful–Ugly	−0.217**	0.000
3 Unique–Ordinary	−0.104**	0.000	3 Unique–Ordinary	−0.134**	0.000
4 Calm–Exciting	0.088**	0.000	4 Happy–Sad	0.056*	0.011
5 Professional–Unprofessional	−0.058**	0.007	5 Slow–Fast	0.066**	0.000
6 Quiet–Loud	−0.055**	0.008	6 Realistic–Unrealistic	−0.047**	0.007
7 Passive–Active	0.053*	0.011	7 Hard–Soft	0.058**	0.002
8 Happy–Sad	0.038	0.056	8 Professional–Unprofessional	−0.062**	0.008
9 Realistic–Unrealistic	−0.050**	0.004	9 Colorful–Colorless	0.051*	0.016
10 Concrete–Abstract	0.042*	0.021			
11 Hard–Soft	0.053**	0.004			
12 Feminine–Masculine	0.043*	0.021			

* $p < 0.05$.
** $p < 0.01$.

to Tables 9 and 10 for results.

The stepwise regressions that were performed without the adjective pairs Beautiful–Ugly and Good–Bad inserted 15 to 11 adjective pairs in the final models per dependent (overall evaluation: 15, willingness to click: 13, willingness to download: 14, and willingness to purchase: 11). All variables in the models were statistically significant at the 0.05 or 0.01 level. The findings repeat our notion regarding the adjective pairs Beautiful–Ugly and Good–Bad, namely, that several underlying significant effects are revealed without these two adjective pairs with the highest VIF scores (Table 5) in the model.

The stepwise regression results indicate that regarding the relationship between consumer perceptions of app icons and their *overall evaluation*, the following ends of the semantic differentials positively predicted their grade (Table 7): *good, beautiful, active, professional,*

Table 9
Overall evaluation and the willingness to click excluding Beautiful–Ugly and Good–Bad (stepwise).

Step #	Beta Evaluation (R^2 = 0.566)	p	Step #	Beta Click (R^2 = 0.387)	p
1 Professional–Unprofessional	−0.293**	0.000	1 Expensive–Cheap	−0.170**	0.000
2 Happy–Sad	−0.096**	0.000	2 Professional–Unprofessional	−0.203**	0.000
3 Expensive–Cheap	−0.189**	0.000	3 Unique–Ordinary	−0.147**	0.000
4 Passive–Active	0.084**	0.000	4 Happy–Sad	−0.075**	0.000
5 Unique–Ordinary	−0.104**	0.000	5 Passive–Active	0.127**	0.000
6 Relaxed–Stiff	−0.117**	0.000	6 Realistic–Unrealistic	−0.089**	0.000
7 Strong–Weak	−0.148**	0.000	7 Quiet–Loud	−0.128**	0.000
8 Old–Young	0.063**	0.000	8 Strong–Weak	−0.126**	0.000
9 Calm–Exciting	0.099**	0.000	9 Calm–Exciting	0.107**	0.000
10 Quiet–Loud	−0.085**	0.000	10 Relaxed–Stiff	−0.067**	0.002
11 Hard–Soft	0.070**	0.000	11 Complex–Simple	0.040*	0.031
12 Feminine–Masculine	0.059**	0.001	12 Hard–Soft	0.065**	0.003
13 Realistic–Unrealistic	−0.046**	0.002	13 Two–Three-dimensional	−0.047*	0.015
14 Colorful–Colorless	−0.043*	0.021			
15 Complex–Simple	0.032*	0.038			

* $p < 0.05$.
** $p < 0.01$.

Table 10
The willingness to download and purchase excluding Beautiful–Ugly and Good–Bad (stepwise).

Step #	Beta Download (R^2 = 0.378)	p	Step #	Beta Purchase (R^2 = 0.304)	p
1 Expensive–Cheap	−0.179**	0.000	1 Expensive–Cheap	−0.153**	0.000
2 Professional–Unprofessional	−0.210**	0.000	2 Professional–Unprofessional	−0.197**	0.000
3 Relaxed–Stiff	−0.080**	0.000	3 Unique–Ordinary	−0.137**	0.000
4 Unique–Ordinary	−0.125**	0.000	4 Relaxed–Stiff	−0.095**	0.000
5 Calm–Exciting	0.119**	0.000	5 Realistic–Unrealistic	−0.099**	0.000
6 Realistic–Unrealistic	−0.100**	0.000	6 Strong–Weak	−0.104**	0.000
7 Quiet–Loud	−0.117**	0.000	7 Hard–Soft	0.076**	0.001
8 Passive–Active	0.080**	.001	8 Slow–Fast	0.086**	0.000
9 Strong–Weak	−0.106**	0.000	9 Quiet–Loud	−0.116**	0.000
10 Hard–Soft	0.059**	.007	10 Calm–Exciting	0.082**	0.001
11 Complex–Simple	0.054**	.003	11 Concrete–Abstract	0.042*	0.038
12 Happy–Sad	−0.055*	.011			
13 Concrete–Abstract	0.043*	.027			
14 Old–Young	0.042*	.030			

* $p < 0.05$.
** $p < 0.01$.

unique, exciting, young, masculine, relaxed, strong, soft and *three-dimensional*. Without the adjective pairs Beautiful–Ugly and Good–Bad, the following adjectives also revealed to be positive predictors of the overall evaluation (Table 9): *happy, expensive, quiet, realistic, colorful,* and *simple*.

The following ends of the semantic differentials positively predicted *the willingness to click app icons* (Table 7): *good, beautiful, unique, active, colorless, exciting, quiet, masculine, two-dimensional, realistic* and *soft*. Without the adjective pairs Beautiful–Ugly and Good–Bad, the following adjectives also revealed to be positive predictors of the willingness to click (Table 9): *expensive, professional, happy, strong, relaxed, and simple.*

The following ends of the semantic differentials positively predicted *the willingness to download the imagined app that the icon belongs to* (Table 8): *good, beautiful, unique, exciting, professional, quiet, active, realistic, abstract, soft* and *masculine*. Without the adjective pairs Beautiful–Ugly and Good–Bad, the following adjectives also revealed to be positive predictors of the willingness to download (Table 10): *expensive, relaxed, strong, simple, happy,* and *young.*

The following ends of the semantic differentials positively predicted *the willingness to purchase the imagined app that the icon belongs to* (Table 8): *good, beautiful, unique, sad, fast, realistic, soft, professional* and *colorless*. Without the adjective pairs Beautiful–Ugly and Good–Bad, the following adjectives also revealed to be positive predictors of the willingness to purchase (Table 10): *expensive, relaxed, strong, quiet, exciting*, and *abstract.*

Lastly, we ran the MLRA models with the variable on how many mobile games participants download per week as a control variable to investigate systematic effect on rating. Section 3.1 (Table 1) stated that the majority of participants (75.4%) did not download any game apps on a weekly basis. Including this variable in the analyses did not alter the ratings in a significant manner. The number of game apps downloaded had a statistically significant effect ($\beta = -0.034$, $p < 0.01$) in the *overall evaluation* of the icon, but as the effect is quite small, it can be considered irrelevant in these results.

5. Discussion

Using semantic differential scale (22 adjective pairs), this study investigated the relationship between consumer perceptions of app icons and icon successfulness, measured by 1) overall evaluation of the icon, 2) willingness to click the icon, 3) willingness to download the imagined app and, 4) willingness to purchase the app. The study design was a vignette study, in which participants (n = 569) were assigned to evaluate 4 randomized icons from a total of pre-selected 68 game app icons across 4 categories (concrete, abstract, character and text). This resulted in a total of 2276 individual icon evaluations. The goal was to discover aesthetic qualities that are likely to predict consumer behavior related to clicking on app icons as well as downloading and purchasing apps on app stores.

A clear pattern was displayed by the ratings of the MLRA including all 22 adjective pairs (Table 4) in that the likelihood to icon successfulness can be predicted by the following set of adjectives: *beautiful* (vs. *ugly*), *good* (vs. *bad*), and *unique* (vs. *ordinary*). Icons that were associated with these adjectives projected a positive overall evaluation and willingness to click the app icon as well as download and purchase the imagined app. The polar opposite of these adjectives on the semantic scale has an equally negative effect on the aspects of icon successfulness.

The adjectives “beautiful” and “good” were statistically significant in all cases, which was to be expected. As the adjectives are of general nature, they may reflect more of an overall estimate of aesthetic quality of an icon which poses a potential limitation that should be considered in future studies. Adjective pairs related to aesthetic qualities, such as Colorful–Colorless, Realistic–Unrealistic and Two-dimensional–Three-dimensional, are perhaps more specific in nature and thus express more variation in the ratings seen on Table 4. Nevertheless, this insight is valuable as it contrasts prior results (Shaikh, 2009) on onscreen typeface design and usage that was implemented as a basis for our experiment. Shaikh's (2009) results indicate that not all typefaces for online content should convey beauty, particularly if it is not consistent with the meaning and context of the text. The findings in this study indicate that in app icon successfulness, beauty is an important factor in all cases regardless of the context.

Prior research emphasizes the importance of icon uniqueness related to task performance and user preference (Arend et al., 1987; Creusen and Schoormans, 2005; Creusen et al., 2010; Dewar, 1999; Goonetilleke et al., 2001; Huang et al., 2002). The results in this study support this notion as the adjective “unique” is statistically significant in each of the four variables in the MLRA including all 22 adjective pairs (Table 4) along with the post-hoc MLRA that were performed without the adjective pairs Beautiful–Ugly and Good–Bad (Table 6). Icon memorability is suggested as one of the key design elements for app icons by the developer guides of leading app stores,[6,7] which is likely due to the large mass of app icon material available for consumers on app stores. Uniqueness helps app icons stand out in a display

[6] Apple Developers, “App icon,” https://developer.apple.com/ios/human-interface-guidelines/icons-and-images/app-icon/ (Accessed January 30, 2018).
[7] Android Developers, “Iconography,” http://www.androiddocs.com/design/style/iconography.html (Accessed January 30, 2018).

of many other offerings. Hence, on the basis of the ratings in our analyses (Tables 4 and 6), we suggest that app icons need to be distinguishable to successfully attract consumers. Evidently, icon uniqueness is a combination of multiple features. However notably, a comparison between the four icon categories (abstract, concrete, character and text) in this study indicate that abstract icons were perceived as unique more often than icons from the other categories. Thus, abstract elements may enhance perceived icon uniqueness.

The post-hoc MLRA that were performed without the adjective pairs Beautiful–Ugly and Good–Bad (Table 6) exposed other significant effects in addition to icon uniqueness, that may perhaps explicate icon successfulness on a more detailed level. The results in Table 6 indicate that the likelihood to a higher overall evaluation as well as clicking, downloading and purchasing can be predicted by the following adjectives: *professional* (vs. *unprofessional*), *expensive* (vs. *cheap*), *soft* (vs. *hard*), *strong* (vs. *weak*), *relaxed* (vs. *stiff*), *realistic* (vs. *unrealistic*), *exciting* (vs. *calm*) and *quiet* (vs. *loud*).

Product presentation has been shown to affect consumer perceptions, meaning that if the presentation conveys high quality, consumers also perceive the product to be of high quality, and vice versa in relationship to low quality presentation (Silayoi and Speece, 2004). This way, the representation can be favorable or unfavorable to the content. The evidence in this study suggests that this theory may apply to app icons, as both the adjectives “professional” and “expensive” convey high quality. From the pool of the 68 game app icons used in this study, these adjectives are associated with such aesthetic app icon qualities as e.g. rounded corners, use of color gradient, shading and highlighting.

Prior research (Burgers et al., 2016) has established a connection with positive consumer attitudes and the use of visual metaphors in app icons. In consumer research, the use of metaphors has been shown to enhance appreciation of a product, because it is much like solving a puzzle, which rewards the consumer and thus inspires positive evaluations (Phillips and McQuarrie, 2009). For example, the product attribute of softness can be metaphorically represented by feathers or kittens, whereas strength can be represented by an image of a lion (Fenko et al., 2018). The statistical significance of the adjectives “soft” and “strong” highlights these prior observations. In this study, “soft” which is the opposite of “hard”, is associated with such aesthetic app icon qualities as e.g. desaturated colors and objects depicted in the icon that are perceived as delicate, e.g. animal fur. On the other hand, “strong” which is the opposite of “weak”, is associated with bold colors and hard-rendered surfaces, as well as objects depicted in the icon that are perceived powerful, e.g. a flying saucer or a hammer. It is believed that positive emotion between consumer and product brings extra value and increases the possibility of purchase (Cho and Lee, 2005). Furthermore, positive impressions have been considered as an important part of consumer perception (Yun et al., 2003). The statistical significance of the adjectives “relaxed” and “quiet” emphasize this observation as they are emotionally engaging qualities that can be perceived positive. Prior results have shown that vivid, highly saturated colors are perceived as exciting by consumers (Labrecque and Milne, 2011). The adjective “exciting”, which is the opposite of “calm” in this study, supports these findings as the icons that received high ratings for the adjective “exciting” express bold colors, similar to the icons perceived as “strong”. The icons perceived as “exciting” also highly correlate with the stimulus depicted in the icon, e.g. the face of an angry dragon or riding a motorcycle.

Icon concreteness is the extent to which it depicts real objects (Isherwood et al., 2007), whereas icon abstractness tends to have less obvious connections with real objects (McDougall et al., 1999). In this study, the positive ratings for the adjective “realistic” may be correlated to icon concreteness, which has been proven a significant predictor in icon effectiveness in prior studies (Hou and Ho, 2013; McDougall and Reppa, 2008; McDougall et al., 1999, 2000; Moyes and Jordan, 1993; Rogers and Oborne, 1987). However, the adjective “concrete” was found to be of no significant effect in the results (Tables 4 and 6), which justifies further research on the relationship of concreteness and realism within the genre of app icons.

The main observation of the results is not only the similarities that increase our insight into this topic, but also the differences in conjunction with the recurrence of statistically significant variables that may explicate consumer perceptions of app icons on a more detailed level. In spite of the findings in the MLRA that were performed without the adjective pairs Beautiful–Ugly and Good–Bad (Table 6), it is important to note that both “beautiful” and “good” are significant in predicting consumer interaction with app icons.

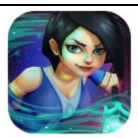
The findings in this study exposed gaps in prior icon design theories, which they did not replicate to the following extent. From the perspective of previous literature on effective icon design, the statistical insignificance of the adjective pairs Concrete–Abstract and Complex–Simple was unexpected. It has been widely speculated that the Concrete–Abstract (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, 2000; Moyes and Jordan, 1993; Rogers and Oborne, 1987) and Complex–Simple (e.g. Choi and Lee, 2012; Goonetilleke et al., 2001; McDougall and Reppa, 2008, 2013; McDougall et al., 2016) relationship predicts icon successfulness. However, the results of our experiment contrast this statement as neither of the adjective pairs was statistically significant in the first MLRA (Table 4). In the second MLRA (Table 6), the adjective pair Complex–Simple was only marginally statistically significant in two of the variables (i.e. overall evaluation and downloading). This calls for more research on app icons, as the Concrete–Abstract and Complex–Simple theories were initially established within other icon genres and have not yet been explored further in relationship to app icons.

5.1. Practical implications

Before setting practical implications, it is essential to understand the scope of the empirical study and the scope of possible guidelines that can be drawn. The broad objective of the study was to increase the understanding of how people's aesthetic perceptions of graphical user interface elements affect people's willingness to interact with those elements. In this study, we selected game app icons as the operationalization and/or case of graphical user interface element, and self-reported overall evaluation as well as willingness to click (the icon), download and purchase the app related to the icon as operationalization of GUI element successfulness. It should be noted that this study did not measure or investigate the relationship between graphical features of the icons and aesthetic perception. The study investigated relationships between the aesthetic perception and willingness to interact. Therefore, the study is unable to directly or reliably inform how graphical user interface elements should be designed in terms of their features, rather it can inform what kinds of aesthetic perceptions graphical user interface elements (i.e. icons) should be brought to evoke. Hence, the recommendations related to graphical features herein are based on qualitative assessment of the mean scores of different adjectives in icon ratings. The results and guidelines are directly applicable to the context of mobile (game) app icons, and with some hesitation, all icons. The results could also be applied all the way up to discussing best practices related to any graphical user interface elements but naturally with increased caution as the external validity diminishes the more general the context in which the knowledge from the results is applied in. Naturally, more similar research is needed across categories of GUIs to enforce and enrich the normative knowledge surrounding the area. Practical implications directly following the empirical results of the study are listed in the following.

Design implication 1: First and foremost, the results unsurprisingly suggest that to increase consumer interaction in terms of app icon successfulness (i.e. overall evaluation, willingness to click an icon as well as download and purchase the imagined app behind the icon), the

Table 11
Top 6 icons with the highest score in overall evaluation, willingness to click the icon, as well as download and purchase the imagined app on a seven-point scale (1 = lowest and 7 = highest).

	Overall evaluation		Willingness to click		Willingness to download		Willingness to purchase	
#	Icon	Mean	Icon	Mean	Icon	Mean	Icon	Mean
1		4.77		4.22		4.00		3.68
2		4.52		4.21		3.95		3.63
3		4.51		4.19		3.85		3.58
4		4.50		4.09		3.81		3.50
5		4.31		4.05		3.77		3.48
6		4.24		3.98		3.77		3.40

app icon should be perceived as high quality which is indicated by the results where the following perceptions predicted app icon successfulness across the board (Tables 4 and 6): *beautiful, good, professional*, and *expensive*. All these adjectives can be associated with general high quality. If cursively investigating the icons that score high on these perceptions, they seem to share some of the following features (Appendix A): transparent parts on the outer layers, color gradients, shading and highlighting as well as high graphical fidelity.

Design implication 2: Separately from perceptions related to high quality (implication 1), *uniqueness* was another strong predictor of icon successfulness (Tables 4 and 6). Therefore, a successful app icon should be unique and memorable to stand out from the app store masses where there is a lot of icon material present. If cursively investigating the icons that score high or low on the *uniqueness–ordinariness* continuum, they seem to share some of the following features (Appendix A): 1) icons rated as *unique* more commonly featured asymmetric and abstract shapes, use of various ends of the color spectrum as well as elements rarely encountered in daily life (e.g. a voodoo doll); 2) icons rated as *ordinary* broadly portrayed concrete, static shapes as well as objects typical to daily life (e.g. a house or a book).

Design implication 3: Beyond all perceptions that predicted all other factors of icon successfulness, *sadness* ($\beta = 0.059$, $p < 0.05$) and *fastness* ($\beta = 0.066$, $p < 0.01$) weakly predicted willingness to purchase the imagined app behind the icon (Tables 4 and 6). If cursively investigating the icons that score high on these factors, they seem to share some of the following features (Appendix A): 1) icons rated as *sad* were generally dominated by a desaturated or dark color palette (e.g. shades of grey or pure black), and they depicted elements that can be perceived as unpleasant; 2) icons rated as *fast* illustrated actions or objects that are typically associated with being rapid (e.g. a motorcycle or an airplane). A related observation is that icons with high scores of perceptions for things that are generally regarded as positive do not necessarily lead to higher icon successfulness. The indication that sadness predicts the willingness to purchase an app behind the icon is one example of this. Moreover, high overall evaluation score does not automatically lead to a high score in the willingness to click the icon, nor in the willingness to download or purchase the imagined app. Thus, it can be argued that app icons should incorporate more than one of the design implications in order to enhance the likelihood to consumer interaction.

Purely as illustrative examples, Table 11 introduces icons with the highest scores in overall evaluation of the icon design, the willingness to click the icon, as well as the willingness to download and purchase the imagined app. However, we wish to note again that this study did not study the relationship between icon features and successfulness per se. Therefore, any relationship between icon feature and success should be regarded as background data augmenting the focus of the study that was on the relationship between perception and successfulness.

5.2. Limitations and future research

App icon design is a complex matter with room for further investigation. This study was one of the first attempts to understand consumer perceptions of app icon successfulness by utilizing game app icons as data collection material. Moreover, this study attempted to rule out non-significant adjectives to aid future research on this topic. Nevertheless, this research has several limitations.

As is natural to any study, some compromises have to be made with

regards to study design as it is impossible to include the entire relevant phenomenon in the scope of a single study. In this study, as having all possible icons or icon categories as stimulus material, we selected one larger domain of iconography. We decided to select game app icons as the stimulus material of the study as 1) mobile game app icons are internally a homogenous category of graphical elements in the sense that they all share the same size, same possible color space and thus should eliminate unforeseen variability, 2) they are familiar to people from before-hand and participants can more effortlessly imagine countering such icons in their normal life, 3) game icons exhibit perhaps more heterogeneity in possible styles compared with icons related to utilitarian software/apps, and therefore, game apps may afford greater external validity and/or generalizability across icons, 4) currently game apps represent a hugely timely phenomenon as game apps are the most popular app category globally by several statistics.[8,9,10]

Icons of new game apps were chosen for this study to eliminate the chance of app and icon familiarity and thus, systematic bias. Hence, the set of icons in this study may not represent the icons of top grossing game apps which consumers more commonly face on app stores. Therefore, consumer perceptions may have been affected by the comparison of top grossing game apps and their icons. This should not be regarded solely a limitation, as the sample of icons in this study represents the majority of game icon material on app stores and may thus give a more realistic perspective on consumer perceptions. However, this might contribute to the fact that the icons used in this study received ratings that can be perceived negative, e.g. ugly and cheap. The sample in this study is a nonprobability convenience sample, therefore it is not necessarily representative of all app store users.

Participants were uninformed about the purpose of the apps behind the icons in the experiment, as this could affect the results. Knowledge of the app may pose a risk of bias in user perceptions, thus the choice was made to eliminate possible confounding variables influencing the main objective of the study. This was further controlled by selecting new game app icons for the experiment that were not widely known. Therefore, this study may be limited in terms of external validity in order to maximize internal validity.

The research question of this study was "How does aesthetic perception of an icon lead to icon successfulness" (as measured via likelihood to click, download, purchase and rate it highly). Hence, also our measurement is focused on icon aesthetics. This study did not include other factors aside from aesthetic qualities that contribute to a consumer's willingness to interact with app icons. In this study, we measured participants' self-reported willingness to interact with the icons presented to them. Alternatively, (intended) behavior could have also been tested by having participants actively click or select icons, or in a field experiment in a real app store to track user behavior. Due to the limitations of the measurement related to the dependent variables used in this study, further investigations could pursuit logging behavioral metrics to increase validity of the study. However, in many cases collecting both the perceptual data on UI aesthetics and user behavior simultaneously may prove difficult.

Future research could be expanded in several directions. For one, investigating the Concrete–Abstract and Simple–Complex relationship regarding app icons would be beneficial, because the findings in this study did not support prior literature to that extent. Game app icons were used in this study to maximize internal validity. This introduces 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. Moreover, future research could map out how participants would describe the imagined app behind the icon to see how icon design affects perceptions on the purpose of the app. As this study employed a large-scale quantitative approach, it provides a broad overview of the relationship between consumer perceptions and app icons. To acquire further, in-depth understanding of the topic, a qualitative approach is recommended. The aim of this study was to explore aesthetic qualities that contribute to consumer perceptions of app icon successfulness, yet other possible factors aside from aesthetic qualities could also be studied in the future, e.g. the number and type of downloaded apps and their effect on perceptions of icon successfulness. Finally, differences in perceptions between cultures, genders and personalities would be an interesting approach for future research to aid designing icons that correspond to the needs of different users.

6. Conclusion

This study explored how consumer perception affects app icon successfulness from an aesthetic perspective. Aesthetic appeal is subjective, which is a probable cause for variations in the results. However, the findings show evidence of consensus that proves an empirical relationship on consumer perceptions and icon successfulness. As the data sample in this study is particularly extensive, the results may be regarded as a contribution to overall knowledge. Revealing this phenomenon may be a building block that eventually leads to further theoretical implications around this topic. Furthermore, the design guidelines in this study assist app designers, developers and marketers when creating a key branding asset for app stores. As such, this study has laid the groundwork for future research that aims to understand consumer perceptions of app icons in graphical user interfaces and online storefronts.

Acknowledgments

This work was supported by the Finnish Funding Agency for Technology and Innovation TEKES (40111/14, 40107/14 and 40009/16) and participating partners, as well as Satakunnan korkeakoulusäätiö (http://www.korkeakoulusaatio.fi/etusivu) and its collaborators.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ijhcs.2019.04.004.

References

Arab, F., Malik, Y., Abdulrazak, A., 2013. Evaluation of PhonAge: an adapted smartphone interface for elderly people. In: Proceedings of the 14th IFIP TC 13 International Conference. Springer, Cape Town, South Africa, pp. 547–554. https://doi.org/10.1007/978-3-642-40498-6_44.

Arend, U., Muthig, K.P., Wandmacher, J., 1987. Evidence for global superiority in menu selection by icons. Behav. Inf. Techol. 6, 411–426. https://doi.org/10.1080/01449298708901853.

Ares, G., Piqueras-Fiszman, B., Varela, P., Marco, R.M., López, A.M., Fiszman, S., 2011. Food labels: do consumers perceive what semiotics want to convey? Food Qual. Pref. 22, 689–698. https://doi.org/10.1016/j.foodqual.2011.05.006.

Black, A., 2017. Icons as carriers of information. In: Black, A., Luna, P., Lund, O., Walker, S. (Eds.), Information design: Research and Practice. Routledge, London, pp. 253–269.

Blankenberger, S., Hahn, K., 1991. Effects of icon design on human-computer interaction. Int. J. Man-Mach. Stud. 35, 363–377. https://doi.org/10.1016/S0020-7373(05)80133-6.

Burgers, C., Eden, A., de Jong, R., Buningh, S., 2016. Rousing reviews and instigative images: the impact of online reviews and visual design characteristics on app downloads. Mob. Media Commun. 4, 327–346. https://doi.org/10.1080/15213269.2016.1182030.

Böhmer, M., Krüger, A., 2013. A study on icon arrangement by smartphone users. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, Paris, France, pp. 2137–2146. https://doi.org/10.1145/2470654.2481294.

Chen, CC., 2015. User recognition and preference of app icon stylization design on the smartphone. In: Stephanidis, C. (Ed.), HCI International 2015 - Posters' Extended

[8] Statista, "Most popular Apple App Store categories in September 2018, by share of available apps," https://www.statista.com/statistics/270291/popular-categories-in-the-app-store/ (Accessed February 25, 2019).

[9] Newzoo, "The global games market reaches $99.6 billion in 2016, mobile generating 37%," https://newzoo.com/insights/articles/global-games-market-reaches-99-6-billion-2016-mobile-generating-37/ (Accessed February 25, 2019).

[10] Statista, "Mobile phone gaming penetration in the United States from 2011 to 2020," https://www.statista.com/statistics/234649/percentage-of-us-population-that-play-mobile-games/ (Accessed February 25, 2019).

Abstracts. HCI 2015. Communications in Computer and Information Science 529 Springer, Cham.. https://doi.org/10.1007/978-3-319-21383-5_2.
Cho, H.-S., Lee, J., 2005. Development of a macroscopic model on recent fashion trends on the basis of consumer emotion. Int. J. Consum. Stud. 29, 17–33. https://doi.org/10.1111/j.1470-6431.2005.00370.x.
Choi, J.H., Lee, H.-J., 2012. Facets of simplicity for the smartphone interface: a structural model. Int. J. Hum. Comput. Stud. 70, 129–142. https://doi.org/10.1016/j.ijhcs.2011.09.002.
Creusen, M.E.H., Schoormans, J.P.L., 2005. The different roles of product appearance in consumer choice. J. Prod. Innov. Manage. 22, 63–81. https://doi.org/10.1111/j.0737-6782.2005.00103.x.
Creusen, M.E.H., Veryzer, R.W., Schoormans, J.P.L., 2010. Product value importance and consumer preference for visual complexity and symmetry. Eur. J. Mark. 44, 1437–1452. https://doi.org/10.1108/03090561011062916.
Crossley, L., 2003. Building emotions in design. Design J. 6, 35–41. https://doi.org/10.2752/146069203789355264.
Dewar, R., 1999. Design and evaluation of public information symbols. In: Zwaga, H.J.G., Boersema, T., Hoonhout, H.C.M. (Eds.), Visual Information for Everyday Use. Taylor & Francis, London, pp. 285–303.
Fenko, A., Schifferstein, H.N.J., Hekkert, P., 2010. Shifts in sensory dominance between various stages of user-product interactions. Appl. Ergonomics 41, 34–40. https://doi.org/10.1016/j.apergo.2009.03.007.
Fenko, A., de Vries, R., van Rompay, T., 2018. How strong is your coffee? The influence of visual metaphors and textual claims on consumers' flavor perception and product evaluation. Front. Psychol. 9, 53. https://doi.org/10.3389/fpsyg.2018.00053.
Forsythe, A., Mulhern, G., Sawey, M., 2008. Confounds in pictorial sets: the role of complexity and familiarity in basic-level picture processing. Behav. Res. Methods 40, 116–129. https://doi.org/10.3758/BRM.40.1.116.
García, M., Badre, A.N., Stasko, J.T., 1994. Development and validation of icons varying in their abstractness. Interact. Comput. 6, 191–211. https://doi.org/10.1016/0953-5438(94)90024-8.
Gittins, D., 1986. Icon-based human-computer interaction. Int J. Man-Mach. Stud. 24, 519–543. https://doi.org/10.1016/S0020-7373(86)80007-4.
Goonetilleke, R.S., Shih, H.M., On, H.K., Fritsch, J., 2001. Effects of training and representational characteristics in icon design. Int. J. Hum. Comput. Stud. 55, 741–760. https://doi.org/10.1006/ijhc.2001.0501.
Hou, K.-C., Ho, C.-H., 2013. A preliminary study on aesthetic of apps icon design. In: Proceedings of the 5th International Congress of International Association of Societies of Design Research. Tokyo, Japan.
Horton, W., 1994. The Icon book: Visual symbols For Computing Systems and Documentation. John Wiley & Sons, New York.
Horton, W., 1996. Designing icons and visual symbols. In: Proceedings of the CHI 96 Conference on Human Factors in Computing Systems. Vancouver, Canada. pp. 371–372. https://doi.org/10.1145/257089.257378.
Huang, S.-M., Shieh, K.-K., Chi, C.-F., 2002. Factors affecting the design of computer icons. Int. J. Ind. Ergon. 29, 211–218. https://doi.org/10.1016/S0169-8141(01)00064-6.
Isherwood, S.J., McDougall, S.J.P., Curry, M.B., 2007. Icon identification in context: the changing role of icon characteristics with user experience. Hum. Fact. 49, 465–476. https://doi.org/10.1518/001872007×200102.
Labrecque, L., Milne, G., 2011. Exciting red and competent blue: the importance of color in marketing. J. Acad. Mark. Sci. 40, 711–727. https://doi.org/10.1007/s11747-010-0245-y.
Lin, C.-H., Chen, M., 2018. The icon matters: how design instability affects download intention of mobile apps under prevention and promotion motivations. Electron. Commer. Res. https://doi.org/10.1007/s10660-018-9297-8.
Lodding, K.N., 1983. Iconic interfacing. IEEE Comput. Graph. Appl. 3, 11–20. https://doi.org/10.1109/MCG.1983.262982.
McDougall, S.J.P., Curry, M.B., de Bruijin, O., 1998. Understanding what makes icons effective: how subjective ratings can inform design. In: Hanson, M. (Ed.), Contemporary Ergonomics. Taylor & Francis, London, pp. 285–289.
McDougall, S.J.P., Curry, M.B., de Bruijin, O., 1999. Measuring symbol and icon characteristics: norms for concreteness, complexity, meaningfulness, familiarity, and semantic distance for 239 symbols. Behav. Res. Methods Instrum. Comput. 31, 487–519. https://doi.org/10.3758/BF03200730.
McDougall, S.J.P., de Bruijn, O., Curry, M.B., 2000. Exploring the effects of icon characteristics on user performance: the role of icon concreteness, complexity, and distinctiveness. J. Exp. Psychol. Appl. 6, 291–306. https://doi.org/10.1037/1076-898X.6.4.291.
McDougall, S.J.P., Reppa, I., 2008. Why do I like it? The relationships between icon characteristics, user performance and aesthetic appeal. In: Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting. New York, USA. pp. 1257–1261. https://doi.org/10.1177/154193120805201822.
McDougall, S.J.P., Reppa, I., 2013. Ease of icon processing can predict icon appeal. In: Proceedings of the 15th international conference on Human-Computer Interaction. Las Vegas, USA. pp. 575–584. https://doi.org/10.1007/978-3-642-39232-0_62.
McDougall, S.J.P., Reppa, I., Kulik, J., Taylor, A., 2016. What makes icons appealing? The role of processing fluency in predicting icon appeal in different task contexts. Appl. Ergon. 55, 156–172. https://doi.org/10.1016/j.apergo.2016.02.006.
Montgomery, D.C., Peck, E.A., 1992. Introduction to Linear Regression Analysis, Second ed. John Wiley & Sons, New York.
Moreira, À.V.M., Filho, V.V., Ramalho, G.L., 2014. Understanding mobile game success: a study of features related to acquisition, retention and monetization. J. Interact. Syst. 5, 2–13.
Moyes, J., Jordan, P.W., 1993. Icon design and its effect on guessability, learnability, and experienced user performance. In: Alty, J.D., Diaper, D., Gust, S. (Eds.), People and Computers VIII. Cambridge University Society, Cambridge, pp. 49–59.
Orth, U.R., Malkewitz, K., 2008. Holistic package design and consumer brand impressions. J. Mark. 72, 64–81. https://doi.org/10.1509/jmkg.72.3.64.
Orth, U.R., Malkewitz, K., 2009. Good from far but far from good: the effects of visual fluency on impressions of package design. Adv. Consum. Res. 36, 211–212.
Overby, E., Sabyasachi, M., 2014. Physical and electronic wholesale markets: an empirical analysis of product sorting and market function. J. Manage. Inf. Syst. 31, 11–46. https://doi.org/10.2753/MIS0742-1222310202.
Peirce, C.S., 1932. Volumes I and II: principles of philosophy and elements of logic. In: Hartshorne, C., Weiss, P. (Eds.), Collected Papers of Charles Sanders Peirce. Harvard University Press, London.
Phillips, B.J., 1996. Defining trade characters and their role in American popular culture. J. Pop. Cult. 29, 143–158. https://doi.org/10.1111/j.0022-3840.1996.1438797.x.
Phillips, B.J., McQuarrie, E.F., 2009. Impact of advertising metaphor on consumer belief: delineating the contribution of comparison versus deviation factors. J. Advert. 38, 49–62. https://doi.org/10.2753/JOA0091-3367380104.
Reppa, I., McDougall, S.J.P., 2015. When the going gets tough the beautiful get going: aesthetic appeal facilitates task performance. Psychol. Bull. Rev. 22, 1243–1254. https://doi.org/10.3758/s13423-014-0794-z.
Rogers, Y., Oborne, D.J., 1987. Pictorial communication of abstract verbs in relation to human-computer interaction. Br. J. Psychol. 78, 99–112. https://doi.org/10.1111/j.2044-8295.1987.tb02229.x.
Rundh, B., 2009. Packaging design: creating competitive advantage with product packaging. Br. Food J. 111, 988–1002. https://doi.org/10.1108/00070700910992880.
Salman, Y.B., Kim, Y., Cheng, H., 2010. Senior-friendly icon design for the mobile phone. In: Proceedings of the 6th International Conference on Digital Content, Multimedia Technology and its Applications (IDC 2010). IEEE, Seoul, South Korea, pp. 103–108.
Schifferstein, H.N., Fenko, A., Desmet, P.M., Labbe, D., Martin, N., 2013. Influence of package design on the dynamics of multisensory and emotional food experience. Food Qual. Prefer. 27, 18–25. https://doi.org/10.1016/j.foodqual.2012.06.003.
Schröder, S., Ziefle, M., 2008. Making a completely icon-based menu in mobile devices to become true: a user-centered design approach for its development. In: Proceedings of the 10th international conference on Human computer interaction with mobile devices and services. ACM, Amsterdam, the Netherlands, pp. 137–146.
Shaikh, A.D., 2009. Know your typefaces! Semantic differential presentation of 40 onscreen typefaces. Usab. N. 11, 23–65.
Silayoi, P., Speece, M., 2004. Packaging and purchase decisions: an exploratory study on the impact of involvement level and time pressure. Br. Food J. 106, 607–628. https://doi.org/10.1108/00070700410553602.
Shu, W., Lin, C.-S., 2014. Icon design and game app adoption. In: Proceedings of the 20th Americas Conference on Information Systems. Georgia, USA.
Underwood, R.L., Klein, N.M., Burke, R.R., 2001. Packaging communication: attentional effects of product imagery. J. Prod. Brand Manage. 10, 403–422. https://doi.org/10.1108/10610420110410531.
van Rompay, T.J.L., Pruyn, A.T.H., Tieke, P., 2009. Symbolic meaning integration in design and its influence on product and brand evaluation. Int. J. Des. 3, 19–26.
Wang, M., Li, X., 2017. Effects of the aesthetic design of icons on app downloads: evidence from an android market. Electron. Commer. Res. 17, 83–102. https://doi.org/10.1007/s10660-016-9245-4.
Wiedenbeck, S., 1999. The use of icons and labels in an end user application program: an empirical study of learning and retention. Behav. Inf. Technol. 18, 68–82. https://doi.org/10.1080/014492999119129.
Woolridge, D., Schneider, M., 2011. Your iOS app is your most powerful marketing tool. In: Woolridge, D., Schneider, M. (Eds.), The Business of iPhone and iPad App Development. Apress, Berkeley, pp. 61–96. https://doi.org/10.1007/978-1-4302-3301-5_4.
Yun, M.H., Han, S.H., Hong, S.W., Kim, J., 2003. Incorporating user satisfaction into the look-and-feel of mobile phone design. Ergonomics 46, 1423–1440. https://doi.org/10.1080/00140130310001610919.

Henrietta Jylhä is a researcher and a PhD candidate at the Gamification Group at Tampere University. Her research focuses on visual aspects in interactive environments related to consumer psychology. She has experience in quantitative methods, i.e. extensive international survey studies and online experiments. She also has a degree in game and computer graphics and a strong background in digital arts. Jylhä's current research explores the relationship between consumer perceptions and app icons.

Juho Hamari is a Professor of Gamification (associate and tenure track) and leads the Gamification Group across the Tampere University and University of Turku. Dr. Hamari has authored several seminal articles on games and gamification from the perspectives of consumer behavior, human–computer interaction, and information systems science. His research has been published in a variety of prestigious venues such as *Organization Studies, JASIST, IJIM, Computers in Human Behavior, Internet Research, Electronic Commerce Research and Applications, Simulation & Gaming*, as well as in books published by among others MIT Press.