Applied Ergonomics 85 (2020) 103073

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: http://www.elsevier.com/locate/apergo

Skeuomorphic or flat icons for an efficient visual search by younger and older adults?

Ruoyu Chen, Jincheng Huang, Jia Zhou *

Department of Industrial Engineering, Chongqing University, China, No. 174 Shazheng Street, Shapingba District, Chongqing, 400044, PR China

ARTICLE INFO

Keywords:
Skeuomorphic design
Flat design
Smartphone
Older adults
App icons
Visual search

ABSTRACT

The debate around skeuomorphic and flat designs has been long lasting and inconclusive, in part because of the lack of empirical evidence supporting the superiority of one or the other icon style from the perspectives of function and aesthetics. Therefore, this study investigated whether older and younger users perceive the aesthetics of icon styles in the same manner as designers and which style results in the most efficient visual search. Using an experimental system that we developed, 24 older and 24 younger participants rated and searched application icons belonging to the two styles. The results indicated that there was generally a notable difference between participants' and designers' perceptions of icon design styles, even after training, and that the perceived icon design styles further influenced the visual search time and accuracy of the participants as well as their evaluation of the icons' beauty. The results imply that the younger participants could use the skeuomorphic icons more efficiently than they could use the flat icons and that they had an advantage over older participants in terms of this ability; however, aesthetically they appreciated flat icons more. In contrast, older participants searched skeuomorphic icons more quickly and accurately than they did flat icons, and aesthetically they appreciated skeuomorphic icons more.

1. Introduction

The increasing number of applications installed on smartphones requires users to frequently search for their target application among other distracting applications. On average, 35 applications used to be installed on smartphones (Think with Google, 2016) and this number has risen to 80 (App Annie, 2018). Moreover, more than 97.8 million Internet users aged above 60 in China have installed an average of 28 applications on their smartphones (China Internet Network Information Center, 2019). Users sometimes need to conduct searches, which are usually inefficient and particularly difficult in the case of rarely used applications. Icon positioning/location and text labels are also used to help visual searches, while the number of locations a user can memorize and the space for text labels underneath icons are usually limited. Icon images are intended to capture the user's attention and thus play an essential role in target application searches.

Icon design is currently influenced by the trend of transitioning from a skeuomorphic to a flat design. The distinction between the two styles is based on subtle design features defined by the designers. A skeuomorphic design uses features such as textures, shadows, and highlights to imitate a real-world object to help users understand the interface (Page, 2014), whereas a flat design removes the three-dimensional (3D) design elements and uses simple flat shapes and bold colors to embrace visual minimalism (Moran, 2015a); thus, flat design is perceived as more attractive by younger adults (Moran, 2016). Early flat design attempts appeared in Microsoft's Metro design language and Windows 8 in 2011, and the launch of Apple's iOS 7 in 2013 indeed pushed the trend toward flat design. In recent practice, a combination of skeuomorphic and flat design has been used, a typical example of which is Google's material design language, called "flat 2.0". This design type is almost flat but uses subtle 3D effects (Moran, 2015b). Such subtle design elements (e.g., textures, shadows, and highlights) stressed by designers might not necessarily be noticed and interpreted by users. Examining the gap between users' perception and designers' perception of two design styles could help reflect on the role of design elements.

There is a lack of consensus on the role of skeuomorphic design as opposed to flat design. In our previous study, we found that, as compared to abstract icons and their associated animations, the use of concrete icons and associated animations to represent real-world objects could help novice older adults achieve a better performance and higher

* Corresponding author.
E-mail address: zhoujia07@gmail.com (J. Zhou).

https://doi.org/10.1016/j.apergo.2020.103073
Received 31 October 2018; Received in revised form 29 January 2020; Accepted 2 February 2020
Available online 10 February 2020

satisfaction level (Zhou et al., 2017). In light of the advantage of imitating real-world objects in icon design, this study examined whether skeuomorphic design is more suitable for older adults than flat design. Given that flat design has become the mainstream style and is well received by younger adults, the choice between skeuomorphic design and flat design may depend on the age of the target user.

The objective of this study was to explore the extent to which users could perceive the difference among icons designed in skeuomorphic and flat styles and how the design style would influence visual searches by younger and older adults. The results of this study will help designers find a suitable design style for people of all ages.

2. Related work

The influence of icon design on visual searches has usually been investigated from the perspective of one or two icon characteristics. Common influential icons characteristics include concreteness, semantic distance, visual complexity, meaningfulness, and familiarity (Isherwood et al., 2007; McDougall et al., 1999). Among these, concreteness is related to skeuomorphic design in that it imitates real-world objects. Concrete icons use more pictorial information to represent real objects, materials, or people (McDougall et al., 1999). Both older and younger adults identified (Leung et al., 2011) and interpreted (Schröder and Ziefle, 2008) concrete icons more accurately, and the observed response time of younger adults when searching concrete icons was also shorter (McDougall et al., 2000; Isherwood et al., 2007). However, advantages that depend on user experience may not be enduring (McDougall et al., 2000). More concrete semantic icons that more closely reflect daily life metaphors facilitated an improvement in the performance of older users (Leung et al., 2011). Icon characteristics are frequently intertwined: there exist strong correlations between concreteness, meaningfulness, and familiarity (McDougall et al., 1999), correlations between concreteness and visual complexity (Isherwood et al., 2007), and overlap between familiarity and complexity/concreteness (McDougall and Reppa, 2008). Therefore, designers may be able to devise a "quick--and-dirty" icon design method.

Skeuomorphic and flat design styles provide a possible means of approaching this issue. However, there is no consensus as to whether skeuomorphic or flat design is superior. A skeuomorphic design can use metaphors to facilitate easy understanding of the function and the digital interaction of screen objects (Blackwell, 2006; Saffer, 2005). An empirical study in which 38 older adults aged from 65 to 91 years participated showed that the higher the degree of realism of a "contact/call/camera" icon, the more highly they evaluated its appearance. Furthermore, older adults could understand the function of a metaphoric icon better (Cho et al., 2015). It was argued that a skeuomorphic design, in which the graphical depiction of an object is close to reality, required less mental effort and was more adaptive for older people who had experienced a decline in their abilities and for people suffering dementia (Blaynee et al., 2016; Kreps et al., 2016). As the technology and users change, a different suggestion is that users are now sufficiently well-trained in the use of the new digital environment that metaphors are no longer necessary and the use of complex ornamentation in the depiction of physical objects has become a meaningless product of designers competing with each other (Page, 2014; Bradley, 2013). This reaction has led to a reversal in the trend, with designers now moving from skeuomorphic to flat design.

When flat design emerged, it received considerable criticism from researchers. They stated that it was too minimalist to convey information to users, resulting in low distinctiveness (Nielsen, 2012), poor affordance (Belveal, 2013; Treder, 2013), and reduced discoverability (Benensohn, 2015). A lack of fundamental design elements (Linowski, 2013) led to usability failure (Hornor, 2015). For example, many users did not click the command "Change PC settings" when they attempted to change the screen background color of Microsoft Windows 8, because everything on the page looked flat and the command resembled a text label rather than a clickable element (Nielsen, 2012). Because of weak clues that an element was clickable (Svennerberg, 2012), users needed to look around more to find the interactive element when they attempted to reserve a room on a flat hotel Web page, which resulted in them spending more time and showing more dispersed fixation than they did on the equivalent skeuomorphic webpage of the same site (Moran, 2017). A broader distribution of fixation was observed when users were asked to click the target icon among flat icons as compared to among skeuomorphic icons. Moreover, the findings of another comparative study indicated that younger users needed almost twice as long to search for a target flat icon in a matrix of flat icons as they did to search for a target traditional icon in a matrix of traditional icons, and oculomotor indicators revealed that the flat icon search was more complex and exerted a higher cognitive load (Burmistrov et al., 2015). However, it was found that icon design style did not significantly affect the younger adults' time to the first fixation, the duration of first fixation, and the time from first fixation to clicking a requested icon among fixed design style icons. Flat icons dispersed a user's focus more than skeuomorphic icons did (Spiliotopoulos et al., 2018).

Nevertheless, in terms of user experience, the flat design was considered superior by younger adults. They tended to consider a flat website as more attractive (Moran, 2016). They experienced more positive emotions and better aesthetics from flat icons than older adults did (Backhaus et al., 2018). Forty-five undergraduates ranging in age from 20 to 27 years rated the user experience of a flat car navigation interface more highly than that of a skeuomorphic interface in terms of stylization cognition, emotional cognition, and decorative cognition (Wu et al., 2015). Similarly, participants ranging in age from 20 to 41 years considered a flat website easier to use and gave a higher system usability score to the flat design than to the skeuomorphic design (Spiliotopoulos et al., 2018).

3. Methodology

3.1. Participants

A total of 48 participants took part in the study. Older adults were recruited from a civic school, whose staff randomly contacted people aged above 60 years by telephone. More participants were then recruited through snowball sampling. As a result, 24 older adults (mean age = 71.7 years, SD = 6.97; 11 men and 13 women) participated in the experiment. Younger adults majoring in engineering and liberal arts were recruited through flyers distributed using the bulletin board system (BBS) of Chongqing University. Twenty-four younger adults (mean age = 23.3 years, SD = 1.71; 8 men and 16 women) participated in the experiment.

3.2. Experimental design

3.2.1. Independent variables

The independent variables of the experiment were icon design style and age. Participant age was divided into two categories: a younger and an older group. The age of the members of the older group was greater than 60 years; the members of the younger group were the college students described above. Design style, a within-subject variable, was also divided into two categories: the skeuomorphic and the flat design style. A set of 72 icons, consisting of pairs of icons for one application that differed in design style (shown in Appendix I), was chosen. The set of icons was taken from the icon library of a mobile phone company in China, Smartisan, which officially released a library where each application icon was designed in both the skeuomorphic and the flat style. These icons represented the designers' perception of skeuomorphic and flat designs.

In addition to the interpretation of icons' graphics, text labels next to icons or even inside icons may help users find a target. Icons with text labels located next to them are superior to icons alone or text alone

Fig. 1. Chinese text labels within icons.

Fig. 2. English text labels within icons.

Fig. 3. No text labels within icons.

(Wiedenbeck, 1999), especially for novice users (Haramundanis, 1996; Leung et al., 2011). Given that the space for text labels underneath icons is usually limited, text labels within icons that might influence the visual search performance were controlled for. Because the older participants might not be able to read and write English, the text labels within icons were divided into three types: Chinese text (shown in Fig. 1), English text (shown in Fig. 2), and no text (shown in Fig. 3). The number of icons of each of the three types was equal.

Other variables that might be influential included the participant's visual acuity, reaction time, finger dexterity, and demographic background. The participants' visual acuity indices were measured by using the Tumbling E chart. The average reaction time (in milliseconds) was obtained through the reaction time test software Inquisit. The finger dexterity (measured in seconds) was obtained using a finger flexibility tester (Bd-601 type). Demographic variables including gender, age, education level, and smartphone and tablet usage experience were obtained through a questionnaire.

3.2.2. Dependent variables

Three dependent variables were used in the experiment: visual search time, number of clicks, and perceived beauty of the icon. The first two variables were automatically recorded by the system used in the experiment. The visual search time was measured from the participant's tap on the "start task" button to his or her tap on the target icon. The number of clicks represented the number of taps made by the participants until he or she successfully found the target icon. The perceived beauty of the icon was measured through a question (Appendix II) on the AttracDiff 2 questionnaire developed by Hassenzahl et al. (2003).

3.3. Task

The participants were required to complete two tasks:

(1) *Icon-rating task.* To determine the differences between users' perception of the skeuomorphic and flat icon design style and that of the designers, an icon-rating task was conducted. Participants were instructed to rate the icons by awarding stars on a seven-point (flat–skeuomorphic) scale to each icon displayed on the screen (Fig. 4). A rating of one to three stars indicated that the participant perceived the icon design as flat and a rating of five to seven stars indicated that the participant perceived the icon design as skeuomorphic. A greater number of stars represented a higher degree of correspondence in the respective ranges. A rating of four stars indicated that the participant was uncertain of

Fig. 4. Interface of icon-rating task.
Note: The task description at the top of the screen states: "Please rate icons on the seven-point scale, where one to three stars indicate flat, five to seven stars indicate skeuomorphic, and four stars indicate neutral. In the respective ranges, a greater number of stars implies a higher degree of correspondence."

Fig. 5. Interface of icon search-and-match task.
Note: The task description at the top of the screen is "Click on the icon that you think matches the function 'Sneeze Reading'."

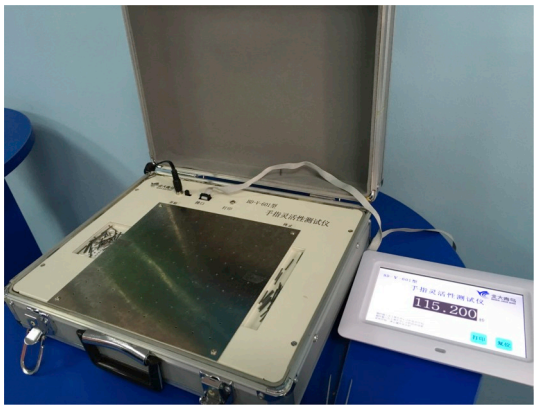


Fig. 6. Finger dexterity tester.

the icon design style. When the 72 icons in the set had been rated, the task was complete.

(2) *Search-and-match task.* To measure the effects of design style on users' visual search performance, a search-and-match task was conducted. The participants were required to click the icon that they considered to match the function name as quickly and accurately as possible (Fig. 5). When participants chose the incorrect icon, its color became gray and it could no longer be clicked. When participants chose the correct icon, the system displayed a message: "Congratulations! Please click 'next' to proceed to the next trial."

3.4. Equipment and materials

The experimental system was run on a notebook computer (ThinkPad S1 Yoga) with the Windows 8.1 operating system and the Google Chrome (version 37.0.2062.120) browser, which was used because its touch screen could avoid the difficulties that older adults experience when using a mouse and keyboard.

Before the experiment, the Tumbling E chart and a color-blindness test book were first used to screen the participants' visual ability. The software Inquisit was used to measure the general reaction time of the participants. The participants tapped the space bar on the computer screen to respond to the observation that red indicators appeared on the software interface; a total of 20 trials were needed to obtain the average response time of the participants. A finger flexibility tester (the Bd-601 type) was used to measure finger dexterity (shown in Fig. 6).

After the experiment, the participants scored the perceived beauty of the flat and skeuomorphic icons on a seven-point Likert scale. In total, six sets of icons were provided (shown in Appendix II), which were grouped by two categories of design styles times three categories of text label types.

3.5. Experimental system

The experimental system was developed using mainly the PHP language. A Bootstrap Star Rating framework was built to support the icon-rating and search-and-match tasks. A series of pilot tests were conducted in which seven subjects participated. The results of each pilot test contributed to the modification of the experimental system, which was refined in seven versions.

For the icon-rating task, a grid of 18 icons was displayed on the computer screen on each webpage, together with instructions at the top

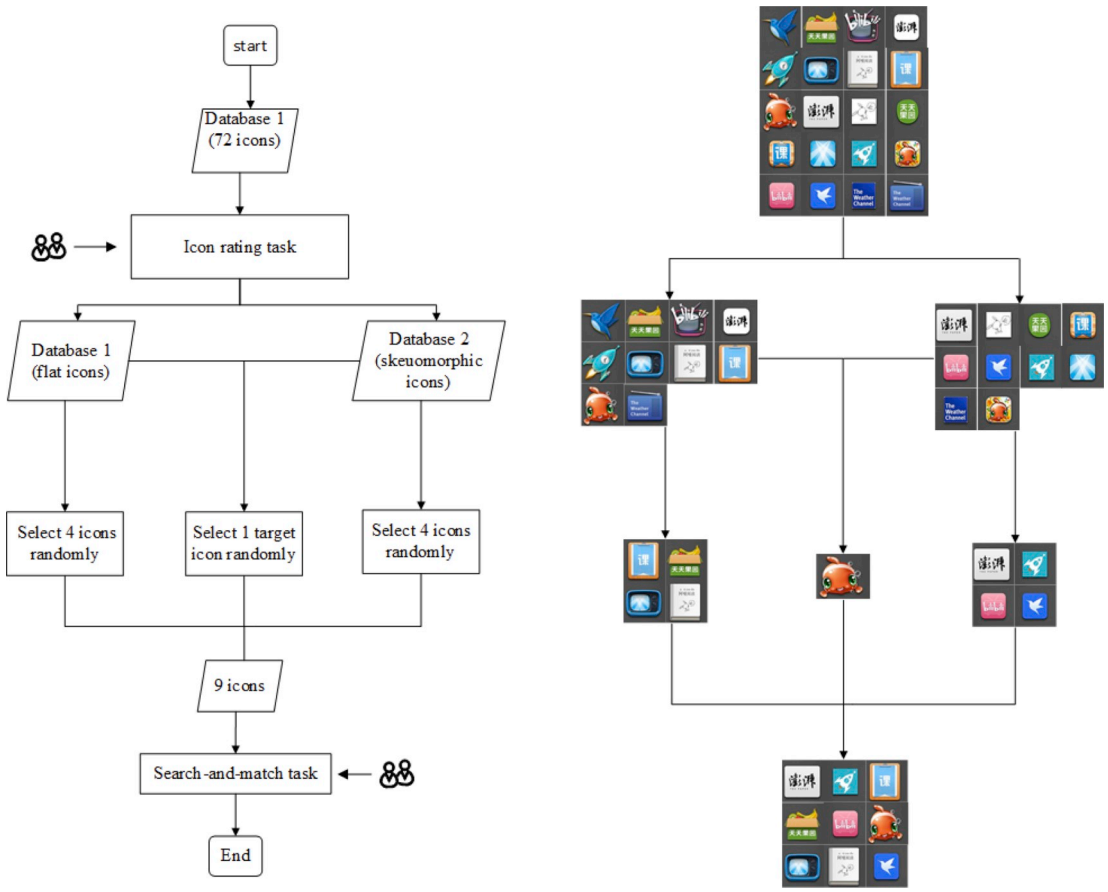


Fig. 7. Experimental system and its sample screens.
Note: The left hand figure shows the icon-selection flow chart used in the search-and-match task in each trial. The right hand figure shows a visual example of the selection of nine icons in every grid displayed on the screen. In the example, the participant divided the standard flat "PengPai" icons into a skeuomorphic and a flat group, which shows his or her view.

of the screen (Fig. 4). The instructions indicated that participants were required to sort the icons by awarding them stars to represent their view. The icons appeared in the sequence shown in Appendix I, and a set of duplicate icons was added to test the consistency of the response of the participants (Set No. 23 and Set No. 35 of icons). Only when the participant had rated each icon on a page could he or she proceed to the next page. The experimental system divided the icons into different icon databases based on the participants' rating results of the 72 icons in the set. Then, the participants performed the search-and-match task.

For the search-and-match task, a grid of nine icons was displayed on the computer screen on each webpage, together with an application function name at the top of the screen (Fig. 5). In each trial, nine icons that consisted of one target icon, four icons that were considered flat, and four icons that were considered skeuomorphic in the icon-rating task were presented randomly at different positions on the grid. A total of 16 trials were conducted: eight trials with flat icon targets and eight trials with skeuomorphic icon targets. In each trial, the nine icons of each grid were selected from the two databases that were established through the icon-rating task (Fig. 7).

3.6. Procedure

The experiment comprised three major steps. First, the participants were briefed and filled out a consent form and a general questionnaire on age, gender, educational background, and technology experience. A color-blindness test book was used to test the participants and those who failed the test were excluded. Then, participants' visual acuity indices, finger dexterity, and reaction time were tested.

Second, pre-test training was conducted. Slides were presented to introduce specific features of the skeuomorphic and flat design styles to participants, including the use of shadows and layers to create depth. These were adapted from the guidelines of Google's material design, a design that is usually considered to be in the middle between skeuomorphic and flat design. To ensure that participants understood the description of the two design styles, they independently rated the design style of 10 icons after training. They could not start the formal test until they reached 70% accuracy.

Third, participants performed the icon-rating task and the search-and-match task independently. Then, a 5-min exploratory interview was conducted. The participants rated the perceived beauty of the six groups of icons (shown in Appendix II) and provided additional comments.

4. Results

4.1. Demographic information

Older participants had a lower education level and less experience than younger participants. Eight older participants had finished junior high school and 16 older participants had finished primary school, while six younger adults were undergraduate students and 18 of them were postgraduate students. Only four older participants used smart phones (mean duration = 1.5 years, SD = 0.60), and older participants never used tablets. In contrast, all younger adults used smart phones (mean duration = 5.5 years, SD = 1.50) and 20 of them used tablets (mean duration = 2.65 years, SD = 1.223). Their mean reaction time, finger

Table 1
Background information on participants.

Variables	Reaction time (ms)		Finger dexterity (s)		Visual acuity		Training accuracy	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Older	882.11	467.40	104.11	24.58	4.64	0.20	79%	0.09
Younger	348.13	54.17	68.80	10.64	4.85	0.18	94%	0.07

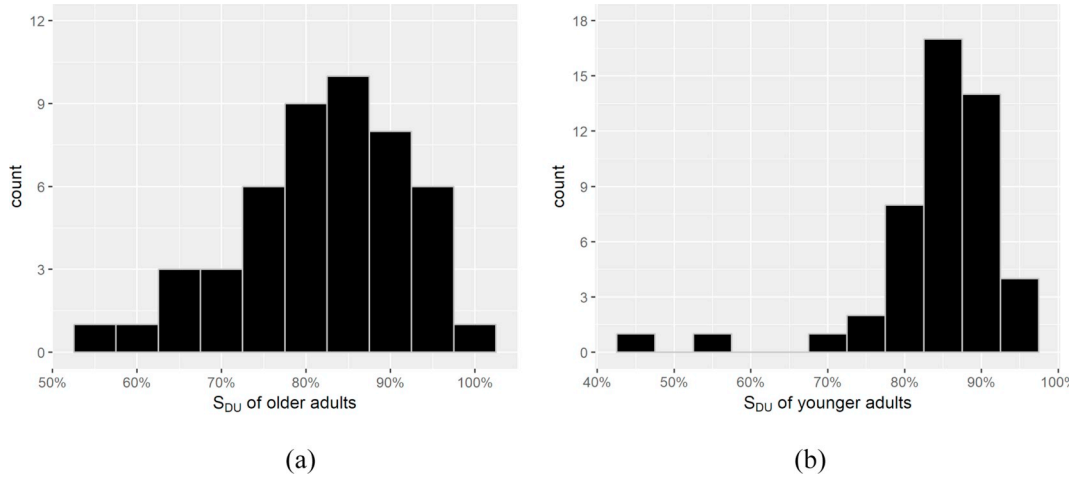


Fig. 8. Distribution of S_{DU} of older adults (a) and younger adults (b).

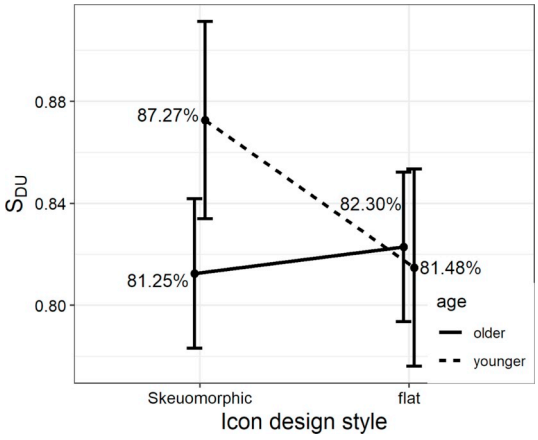


Fig. 9. S_{DU} for participants of different age groups.

dexterity, visual acuity, and training accuracy are shown in Table 1. One older participant with very slight dyserythrochloropsia was not excluded because he still reached 70% accuracy in the training test.

4.2. Perception of icon styles

Statistics for each icon rating task result for each participant were calculated to analyze the extent to which the users' perception of the skeuomorphic and the flat design style of application icons differed from that of the designers. The participants' categorization was counted as agreement with that of the designers only if the participants gave flat icons one to three stars or skeuomorphic icons five to seven stars. The count was computed according to the user and icon to observe the similarity between users' and designers' categorizations (Appendix III).

The measure of similarity is the level of agreement between users and designers (S_{DU}). The agreement is expressed as a percentage, which is calculated by dividing the number of ratings that showed the users and designers to be in agreement by the total number of ratings. For example, of 36 icons categorized by the designer as flat, a user rated 27 icons as flat. In this case, the S_{DU} value is 75%. This measure aimed to quantify the extent to which the participants and designers differed in subjective perception of the icon design style.

The average S_{DU} of older participants was 81.77% and that of younger participants was 84.38%. Because all the participants were trained to reach a 70% similarity level before the test, the results were considered acceptable. Fig. 8 shows the distribution of S_{DU} of the older and younger adults. It can be seen that most younger participants agreed with the designers with a relatively small disparity, while the differences between the older adults and the designers were very significant.

The ANOVA analysis result for the S_{DU} of different age groups and icon design style is shown in Fig. 9. There were no main effects of age ($F_{(1,46)} = 1.270$, $p = 0.266$) and icon design style ($F_{(1,46)} = 1.605$, $p = 0.212$), but a significant interaction existed ($F_{(1,46)} = 4.765$, $p = 0.034$). Bonferroni post hoc tests revealed that the younger participants tended to reach a higher agreement with designers on skeuomorphic icons than on flat icons ($t = 2.147$, $p = 0.0426$), and the agreement on skeuomorphic design was higher for younger than for older participants ($p = 0.032$).

To examine further the extent to which the S_{DU} influenced the performance in the two age groups, a linear regression analysis was conducted. The results showed that, although a higher similarity between

Table 2
Predictors of visual search time.

	Standard coefficients	t	p	R^2 change
Age	0.2313	20.39	<0.001	0.3518
S_{DU}	−10.4403	−2.80	0.00523	0.0074

Table 3
Predictors of number of clicks.

	Standard coefficients	t	p	R^2 change
Age	0.0159	11.07	<0.001	0.1379
S_{DU}	−0.3190	−0.673	0.501	0.0016

Fig. 10. Icons having a younger adult S_I of less than 50%.

users and designers contributed to a more efficient visual search, its contribution was limited, accounting for 0.74% of the total variance in visual search time (Table 2) and 0.16% of the total variance in the number of clicks (Table 3). This implies that 100% similarity is not necessary for an efficient visual search.

Another measure, the similarity between the perceptions of icons (S_I), was calculated by dividing the number of participants who agreed with the designer on the design style of each icon by the total number of participants who rated the icon. For example, 12 out of 24 older users rated as flat a certain icon which was categorized by the designer as flat. In this case, S_I is 50%. This measure aimed to identify whether there were icons on which most participants showed disagreement with designers, which helped find the possible reasons for disagreement by analyzing the features of these icons.

The average S_I of skeuomorphic icons was 84.26% and that of flat icons was 81.89%. A closer view of the icons having a low or high S_I value is presented in Fig. 10. For 10 icons, the S_I for younger participants is lower than 50%, which means that half of the younger participants did not categorize them into the same style as the designers did. The younger participants were confused about the style of three pairs of these icons.

4.3. Performance

The data from 767 trials were analyzed (the data from one trial were not complete and were excluded from further analysis). On average, older participants clicked 1.80 (SD = 1.21) times for the skeuomorphic target icon and 2.15 (SD = 1.33) times for the flat target icon. On average, younger participants clicked 1.11 (SD = 0.50) times for the skeuomorphic target icon and 1.19 (SD = 0.53) times for the flat target icon. Given that visual search time and the number of clicks are significantly correlated ($r = 0.68$, $p < 0.001$), only the results of visual search time are presented in subsequent sections.

On average, older participants spent 14.26 s (SD = 9.69) searching for the skeuomorphic target icon and 17.01 s (SD = 10.42) searching for the flat target icon. On average, younger participants took 2.89 s (SD = 2.52) to search for and match the skeuomorphic target icon and 4.68 s (SD = 4.30) to do so for the flat target icon. To identify influential factors for visual search performance, a multiple linear regression analysis was conducted by the stepwise method. Eleven factors (i.e., age group, icon design style, type of text label, gender, education level, tablet use, smartphone use, reaction time, finger dexterity, visual acuity, and training accuracy) entered the model. The type of text label was recoded as two dummy variables: English text and Chinese text (contrasting with no text). The education level was recoded as three dummy variables, junior high school, undergraduate, and postgraduate (contrasting with primary school). The results (shown in Table 4) indicated that the icon design style, type of text label, education level, finger dexterity, and training accuracy could predict 43.56% of the variance in visual search time ($F_{(8,758)} = 74.9$, $p < 0.001$).

Compared with flat icons, skeuomorphic icons required 16.17% less visual search time for older participants and 38.25% less visual search time for younger participants (shown in Fig. 11a). In line with a previous study on the two design styles, younger participants' needed a longer time when searching for flat icons than when searching for skeuomorphic icons (Burmistrov et al., 2015).

Visual search time also varied across the language of text labels and education levels (Fig. 11b) ($\chi^2_{(6)} = 25.10$, $p = 0.0003$). Icons with Chinese text labels required 28.08% less completion time for participants compared to icons without text labels, which was consistent with previous findings on the advantages of text labels (Haramundanis, 1996; Leung et al., 2011; Wiedenbeck, 1999). However, compared with icons without text labels, icons with English text labels resulted in a 22.07% longer visual search time for participants from junior high school but a 33.97% shorter visual search time for postgraduate participants. What is more, participants with poorer finger dexterity spent longer on the search-and-match task (VIF = 2.11, tolerance statistics = 0.47).

However, it was found that the education level was strongly correlated with age ($r_b = 0.98$, $p < 0.001$). Therefore, to check the role of education level, multiple linear regression analysis was conducted for each age group. The results indicated that education level did not influence the visual search time of younger participants or older participants (shown in Table 5). The older participants' visual search time was significantly affected by icon design style, type of text label, and finger dexterity ($F_{(6,376)} = 10.39$, $p < 0.001$), whereas younger participants' visual search time was just highly related to which type of text label an icon had ($F_{(2,381)} = 10.12$, $p < 0.001$). This implied that education level was no longer influential in explaining the performance difference within each age group. Based on this, an effect of age on visual search performance could not be ruled out.

4.4. Perceived beauty

A multiple linear regression analysis was also conducted for perceived beauty. The education level and icon design style significantly predicted 16.63% of the variance in perceived beauty ($F_{(7,284)} = 9.18$, $p < 0.001$). As shown in Fig. 12a, participants with a lower education level perceived the skeuomorphic icons as more beautiful than flat icons compared to participants with a higher education level (postgraduate vs. junior high, $t = -3.65$, $p < 0.001$; undergraduate vs. junior high, t =

Table 4
Predictors of visual search time.

Variables	B(SE)	Standard Coefficient	t	p	R^2 change
Icon design style (skeuomorphic)	−2.018 (0.524)	−0.105	−3.846	<0.001[a]	0.99%
Type of text label (Chinese text)	−3.184 (0.643)	−0.153	−4.948	<0.001[a]	3.30%
Type of text label (English text)	1.004 (0.632)	0.049	1.589	0.113	
Education level (junior high)	−0.012 (0.808)	0.000	−0.014	0.988	37.52%
Education level (undergraduate)	−14.817 (0.904)	−0.518	−16.388	<0.001[a]	
Education level (postgraduate)	−15.087 (1.069)	−0.745	−14.111	<0.001[a]	
Finger dexterity	0.058 (0.015)	0.157	3.972	<0.001[a]	1.56%
Training accuracy	−6.465 (3.439)	−0.074	−1.880	0.061	0.19%

[a] Coefficient is significant at the 0.001 level.

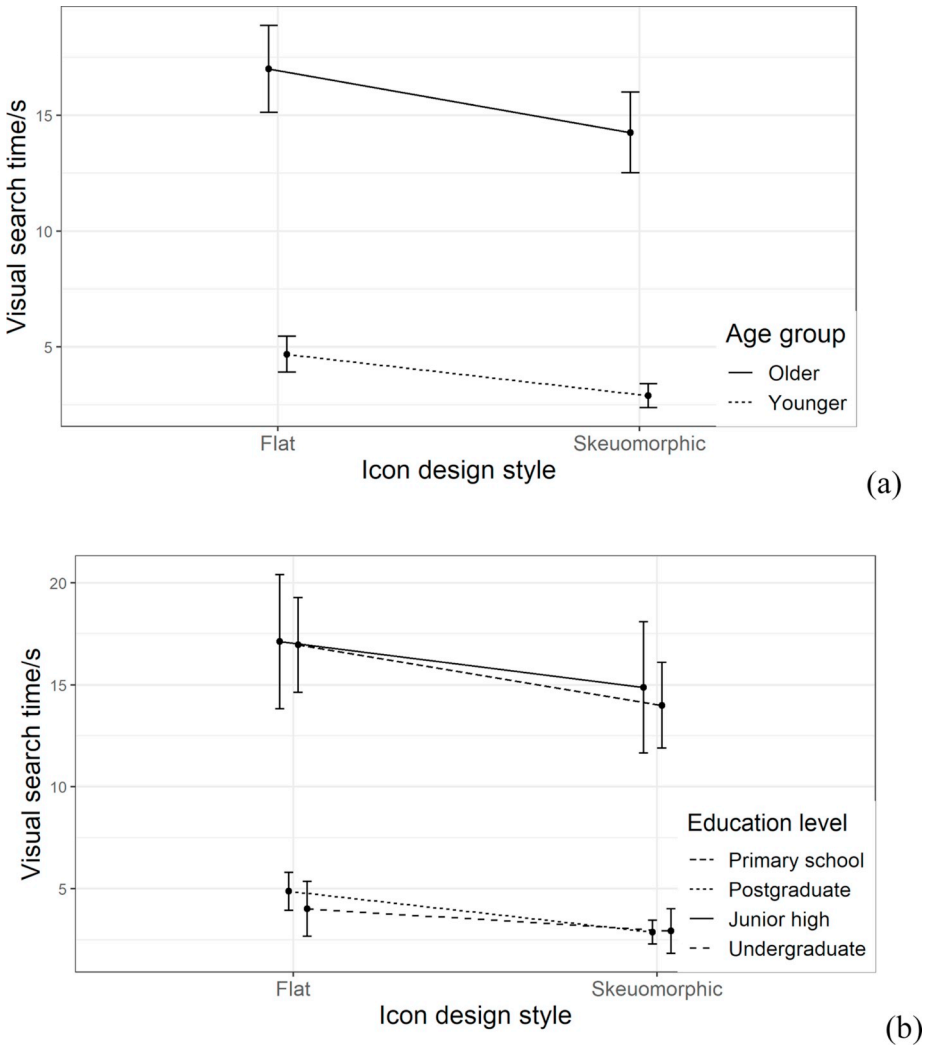


Fig. 11. The influence of design style, age group (a), and education level (b) on visual search time.

Table 5
The results of multiple linear regression analysis for the visual search time of different age groups.

Variables	B(SE)	Standard Coefficient	t	p	R^2 change
Older age group					
Icon design style (skeuomorphic)	−2.644 (0.97)	−0.131	−2.73	0.007[a]	1.59%
Type of text label (Chinese)	−3.839 (1.21)	−0.176	−3.18	0.002[a]	7.64%
Type of text label (English)	3.140 (117)	0.148	2.68	0.008[a]	
Finger dexterity	0.077 (0.02)	−0.186	3.31	0.001[a]	2.87%
Practice accuracy	−10.330 (5.91)	−0.090	−1.75	0.081	0.48%
Gender (male)	1.833 (1.11)	0.090	1.66	0.099	0.28%
Younger age group					
Type of text label (Chinese)	−0.24 (0.06)	−0.322	−3.88	<0.001[b]	4.55%
Type of text label (English)	−0.23 (0.06)	−0.165	−3.77	<0.001[b]	

[a] Coefficient is significant at the 0.01 level.
[b] Coefficient is significant at the 0.001 level.

−2.07, p = 0.04; postgraduate vs. primary, t = −1.199, p = 0.048; undergraduate vs. primary, t = −4.04, p < 0.001). Similarly, as shown in Fig. 12b, the older adults considered the skeuomorphic icons as more beautiful (t = 4.412, p < 0.01) than the flat icons, while the younger adults considered the flat icons as more beautiful (t = 2.215, p = 0.010).

5. General discussion

The essential premise of the discussion on skeuomorphic and flat design is that there is consensus on the categorization of the two design styles. Previous debates about the two styles of icons were based on their categorization by designers; however, designers' categorization is not necessarily the same as that of users. Although training was conducted before the experiment, a nearly 20% difference was found between designers' and users' perceptions of the two icon styles. This implies that, although users could distinguish the majority of skeuomorphic and flat icons created by designers, there is a middle ground where users' and designers' perceptions conflict. This disagreement may be related to the fact that current practices for distinguishing skeuomorphic and flat designs emphasize mainly salient visual features, such as depth and texture (Creager and Gillan, 2016). However, people usually searched for the stimuli not only through salient visual features (e.g., shape, color, and boarder thickness) but also through semantic information or a combination of the two (Tao et al., 2017). Designers should not rely completely on the salient visual features of application icons but should

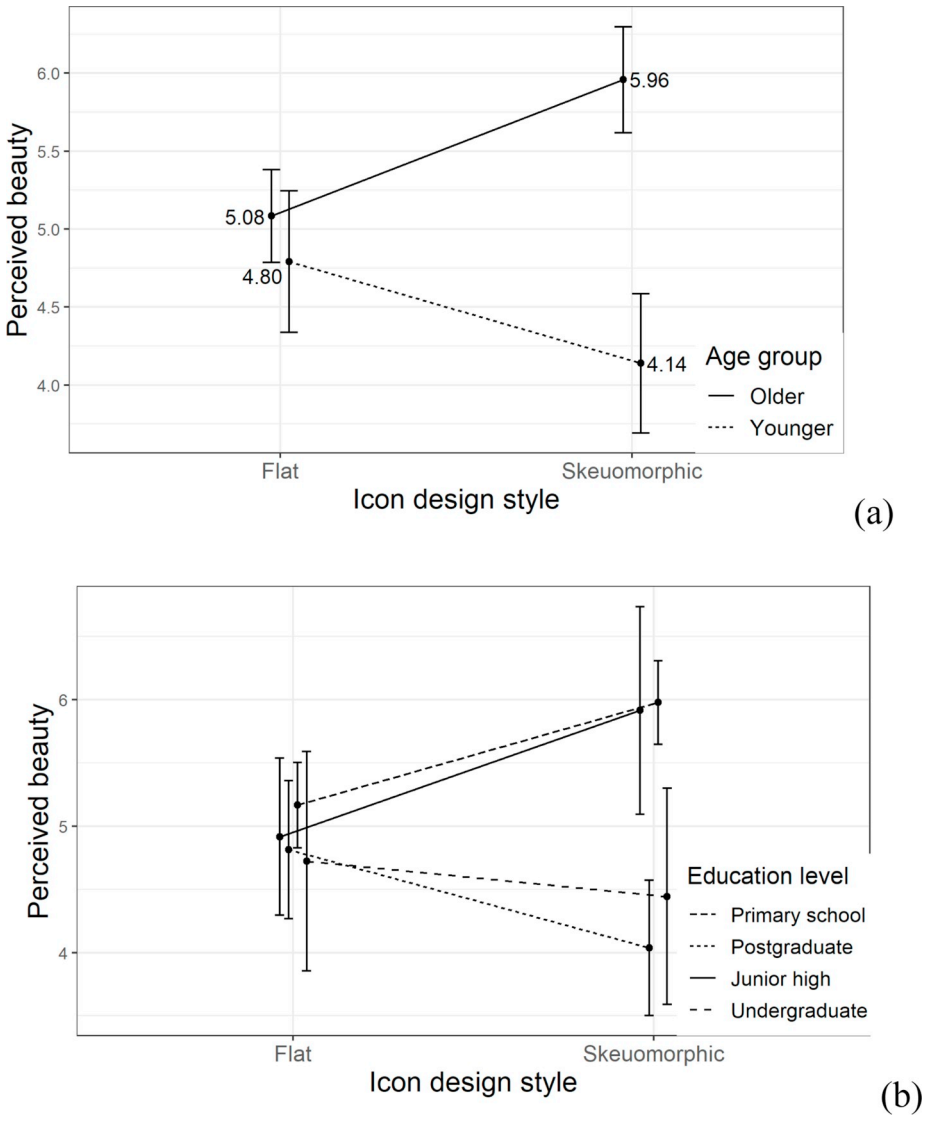


Fig. 12. Average perceived beauty score for each group of education level (a), age (b) and icon design style.

also consider the understandability of application icons' semantic information.

Skeuomorphic application icons helped both older and younger participants to search more quickly and accurately than flat application icons did. The older participants' performance was better for skeuomorphic icons, which confirms the idea derived in a previous experiment among American older adults (Zhou et al., 2017) and is consistent with interview results among Korean older adults (Cho et al., 2015). The finding beyond our expectations is that the performance of the younger adults was also better for skeuomorphic icons. One possible reason is that younger adults had had more experience with skeuomorphic design, which was favored in the initial design of iOS and Microsoft Windows, whereas flat design is relatively new and its simplicity/minimalism could interfere with users' cognitive understanding of icons if not carefully applied (Gu, 2016). Another possible reason is that skeuomorphic icons are more realistic and convey metaphors better. A questionnaire survey among 155 younger adults found that, regardless of their familiarity with icons, they identified skeuomorphic icons more accurately than flat icons (Zhang et al., 2017).

Although skeuomorphic application icons outperformed flat icons in terms of their effectiveness, the younger participants considered the flat application icons more beautiful. The reason for the contrast between the performance and the aesthetic evaluation among younger adults may be that the flat design has become the new "fashion" (Pan and Stolterman, 2015) and is thus easily followed by younger adults. In addition, the flat design adopts minimalism as the key approach (Gu, 2016), which may be consistent with younger adults' desire to adopt a minimalist lifestyle (Weinswig, 2016). This might imply that younger adults might be willing to sacrifice usability for aesthetics under certain conditions. A possible means of balancing usability and aesthetics is to lay subtle salient visual features on a completely flat design, making elements more easily distinguishable and thus providing a medium approach between skeuomorphic and flat design. This design style is termed "flat 2.0" (Creager and Gillan, 2016; Moran, 2015b). Designers should not merely pursue trends of skeuomorphic or flat design styles. Instead, it is more important to consider the initial intention of a design and the reasons for changing it.

Older adults perceived skeuomorphic icons as more aesthetically appealing. However, the advantage of skeuomorphic design over flat design in terms of both performance and aesthetics does not mean that designers should revert to the skeuomorphic design trend. Given that previous studies showed that when users became familiar with icons the advantage of concrete icons over abstract icons was diminished (McDougall et al., 2000; Isherwood et al., 2007) or lasting (Schröder and Ziefle, 2008), whether the advantage of skeuomorphic design would be enduring needs further study.

Besides the icon design style, the type of text label affected the visual search performance. An icon with a Chinese text label was a good choice for native speakers but whether application of an English text label may depend on the target users' education level. Moreover, the participants' performance was affected not only by design features but also by physical functions such as finger dexterity, which were especially significant for older participants. We speculated that the decline in physical ability was a possible reason for differences of performance between the two age groups.

Five limitations of this study should be noted. (1) The perception of icon design styles differed between individual participants. Although in pre-training the distinction between skeuomorphic and flat design was explained according to Google's mobile design standards, participants' understanding may diverge a little from the standard design guidelines. (2) Ideally, the skeuomorphic and flat icons used in experiments should be produced by the key players in the smart phone market; however, the icons of only one mobile phone company were used in this study to present designers' perceptions of the two icon design styles. (3) Although we have conducted linear regression analysis for different age groups and found that education level had no significant effect, the confounded effects of age and education level were difficult to separate because the age and the education level of the participants were not balanced. (4) Furthermore, different devices could even influence a participant's icon search and click behavior. In the experiment, the test system was run on a notebook computer. However, the icons that were displayed on the tablet screen in the experiment are usually displayed on a smart phone. The display on a smart phone and a tablet screen differs in terms of size, color, and scale, and even the interaction method is different. (5) The application icons appeared in groups rather than individually when evaluated for perceived beauty.

Future work may further investigate the patterns of the skeuomorphic and flat icons categorized by users, which would lead to additional understanding of the link between the details of design patterns and visual search performance. In addition, a large sample size and an increased variety of participant demographics (cognition differs by gender, education level, profession, work, culture, and area) are needed. Finally, the use of icons from multiple companies or designers in the industry would produce more reliable results.

6. Conclusions

This study focused on younger and older adults' perceptions of the design styles of application icons, examining the extent of the difference between their perception and that of designers, and investigated the manner in which their perception of skeuomorphic and flat icons influenced their search-and-match performance.

There were notable differences in the icon style perceptions of users and designers. Moreover, older and younger participants performed differently according to whether they perceived the application icons as skeuomorphic or flat. (1) Older participants spent more time and clicked more when searching for icons than younger participants. (2) Skeuomorphic icons contributed to a better performance in terms of visual search time and the number of clicks as compared to flat icons for both age groups. (3) Skeuomorphic icons conformed to older participants' aesthetics, whereas flat icons conformed to younger adults' aesthetics. (4) Both age groups performed better on icons with a Chinese text label than on icons without a text label.

The results of this study have two implications. First, it would be useful to provide skeuomorphic icons, so that older adults can perform better and have a superior user experience, and to add Chinese text identification to icons. Second, the provision of skeuomorphic icons allows younger adults to perform better, while flat icons are more aesthetically pleasing to younger adults.

Acknowledgements

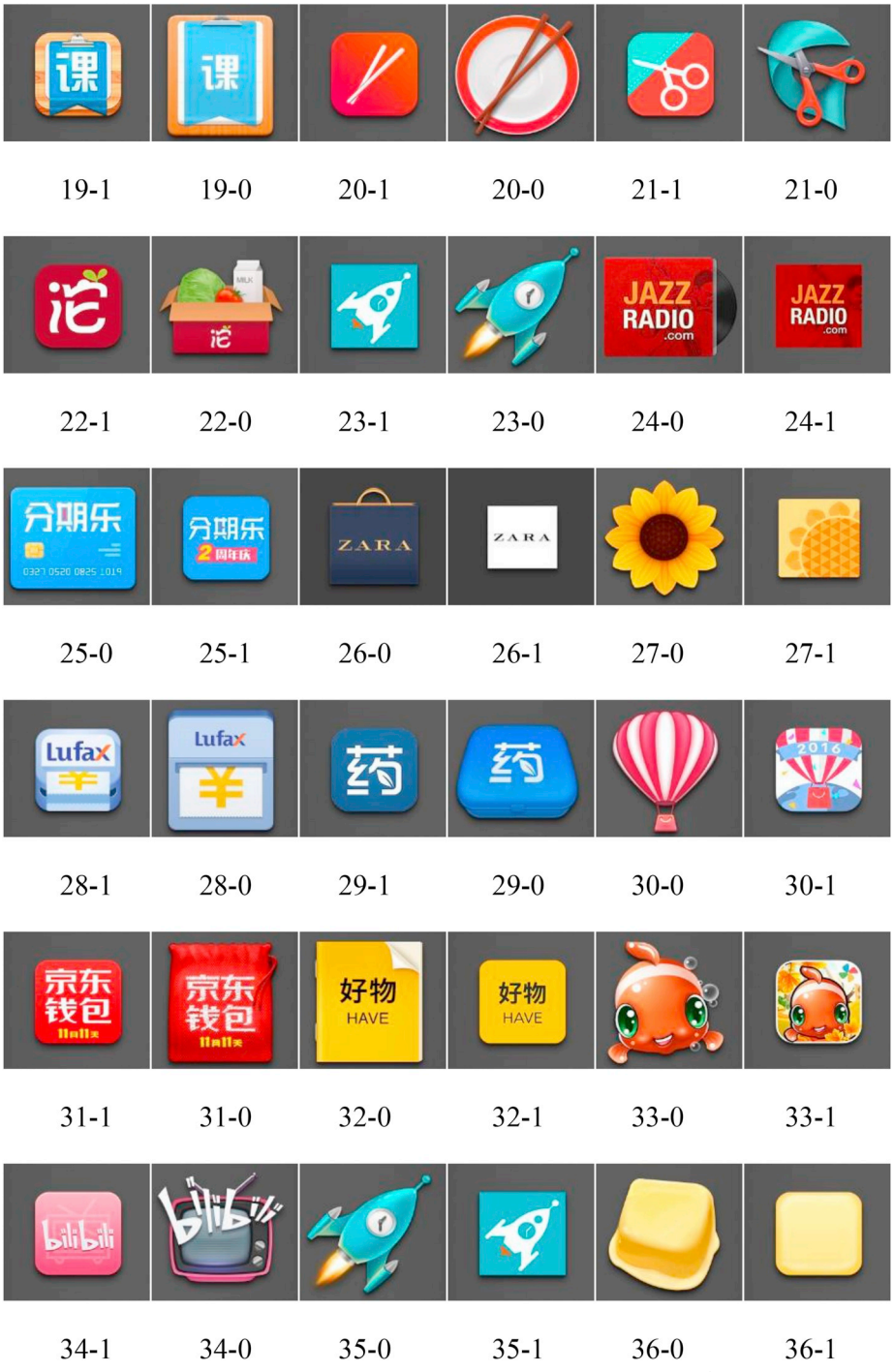
This work was supported by the Chongqing Municipal Natural Science Foundation, under Grant number cstc2016jcyjA0406 and the National Natural Science Foundation of China under Grant number 71661167006.

Appendix I

Icons Used in the Experiment

The first digit under each icon represents the smart phone application, where the same number represents the same app. The second digit represents designer perception of the design style of each icon, where "0" indicates skeuomorphic design style, and "1" indicates flat design style.


DOUYU
DOUYU
澎湃
澎湃
THE PAPER
The Weather Channel
The Weather Channel
阿唯阅读
央视新闻
央视新闻
SAFE
SAFE
Happy Birthday
NCE
NCE
TOPTRIP
Meet you
Meet you
席
席
天天果园
天天果园
1-1
1-0
2-1
2-0
3-0
3-1
4-0
4-1
5-1
5-0
6-1
6-0
7-0
7-1
8-0
8-1
9-1
9-0
10-0
10-1
11-1
11-0
12-0
12-1
13-0
13-1
14-1
14-0
15-0
15-1
16-1
16-0
17-1
17-0
18-0
18-1



. (*continued*).

Appendix II

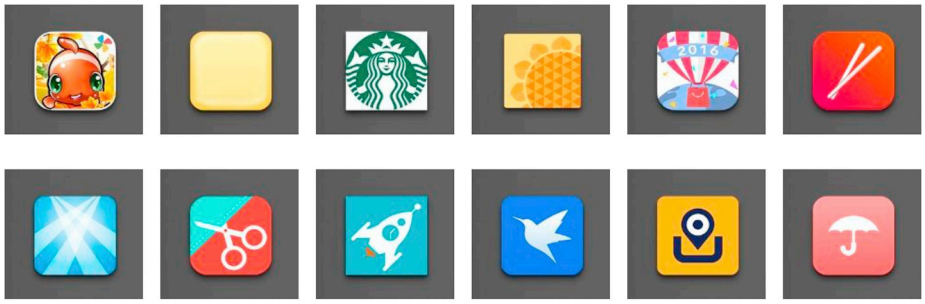
Perceived beauty rating

Please look at icons in each group and rate the extent of agreement or disagreement.

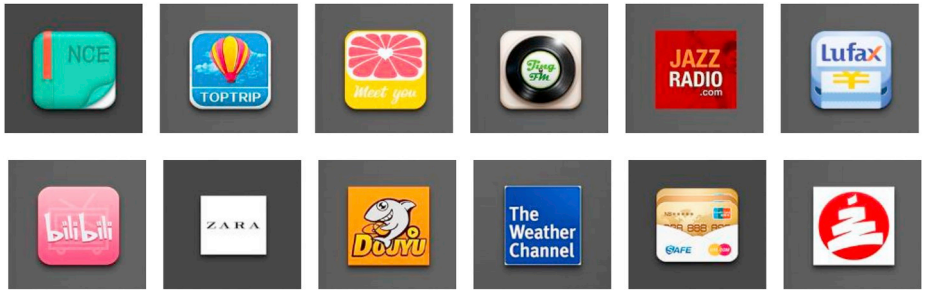
Table II
A 7-point scale of perceived beauty

	Strongly Disagree	Disagree	Slightly Disagree	Neutral	Slightly Agree	Agree	Strongly Agree
I think icons in this group are beautiful.	1	2	3	4	5	6	7

Group 2:



Group 3:



Group 4.

Group 5:

Group 6:

Appendix III

Table III1 Examples of older participants' icon rating results

Icon_id	User_id														
	1	2	3	4	5	6	…	19	20	21	22	23	24	Count	S_I (%) = C/24
1–0	6	6	7	7	7	7	…	6	7	6	6	7	6	24	100.00
2–0	1	6	7	4	4	4	…	3	1	4	4	2	5	6	25.00
3–0	7	6	7	6	7	6	…	7	6	6	6	7	7	23	95.83
4–0	6	6	7	6	7	5	…	5	6	5	3	4	6	15	62.50
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
33–0	6	5	7	7	7	5	…	5	7	6	5	6	6	24	100.00
34–0	6	5	7	6	5	4	…	6	6	5	5	7	7	22	91.67
35–0	6	6	7	6	7	6	…	6	6	6	3	7	6	23	95.83
36–0	7	5	6	7	7	6	…	6	6	6	5	6	6	24	100.00
Count	31	32	36	34	34	30	…	24	28	30	26	30	28	/	/
S_{DU} (%) = C/36	86.1	88.9	100	94.4	94.4	83.3	…	66.7	77.8	83.3	72.2	83.3	77.8	/	/
1–1	6	2	3	3	2	3	…	3	3	3	4	1	4	18	75.00
2–1	1	2	2	1	1	3	…	4	6	2	4	1	4	17	70.83
3–1	1	2	2	3	2	6	…	4	1	3	3	4	3	20	83.33
4–1	1	2	3	2	1	3	…	1	1	3	2	3	3	20	83.33
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
33–1	3	4	2	2	3	3	…	5	3	2	4	1	2	18	75.00
34–1	2	2	3	2	1	4	…	3	2	3	3	2	3	23	95.83
35–1	6	3	3	3	2	3	…	3	3	3	6	3	3	22	91.67
36–1	1	1	2	1	2	2	…	1	1	2	2	2	7	23	95.83
Count	29	35	32	33	34	31	…	28	27	31	26	33	19	/	/
S_{DU} (%) = C/36	80.6	97.2	88.9	91.7	94.4	86.1	…	77.8	75.0	86.1	72.2	91.7	52.8	/	/

Note: A-0 indicates a skeuomorphic icon, A-1 indicates a corresponding flat icon.
C=Count.

Table III
2 Examples of younger participants' icon rating results

Icon_id	User_id						…								
	25	26	27	28	29	30	…	43	44	45	46	47	48	Count	S_I (%) = C/24
1–0	6	6	7	7	7	7	…	7	5	6	4	7	6	23	95.83
2–0	2	3	5	4	1	4	…	5	4	4	4	5	4	4	16.67
3–0	6	7	6	7	7	7	…	7	7	6	5	7	7	24	100.00
4–0	5	6	6	7	4	7	…	6	5	6	4	5	5	21	87.50
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
33–0	7	7	5	7	7	7	…	7	6	7	5	7	6	24	100.00
34–0	7	7	6	7	7	7	…	7	5	7	5	6	6	24	100.00
35–0	7	7	6	7	7	7	…	7	5	6	6	6	7	24	100.00
36–0	5	7	6	7	7	7	…	7	6	6	6	6	6	24	100.00
Count	31	32	34	32	32	32	…	33	30	30	25	30	31	/	/
S_{DU} (%) = C/36	86.1	88.9	94.4	88.9	88.9	88.9	…	91.7	83.3	83.3	69.4	83.3	86.1	/	/
1–1	2	1	3	1	4	4	…	2	3	2	2	1	2	22	91.67
2–1	4	3	3	7	5	4	…	2	4	4	2	5	2	12	50.00
3–1	2	1	2	1	1	4	…	2	1	3	1	1	1	23	95.83
4–1	2	2	3	1	7	3	…	1	2	3	1	3	1	23	95.83
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
33–1	3	3	3	5	4	4	…	4	4	3	2	3	3	16	66.67
34–1	3	2	2	1	2	4	…	3	4	2	2	2	3	21	87.50
35–1	3	1	1	2	3	3	…	1	1	3	2	1	1	23	95.83
36–1	1	1	3	2	2	4	…	2	3	2	3	2	1	22	91.67
Count	28	32	31	29	27	20	…	30	27	29	33	28	31	/	/
S_{DU} (%) = C/36	77.8	88.9	86.1	80.6	75.0	55.6	…	83.3	75.0	80.6	91.7	77.8	86.1	/	/

Note: A-0 indicates a skeuomorphic icon, A-1 indicates a corresponding flat icon.
C=Count.

References

Annie, App, 2018. 2017 retrospective: a monumental year for the app economy. https://www.appannie.com/cn/insights/market-data/app-annie-2017-retrospective/.

Backhaus, N., Trapp, A.K., Thüring, M., 2018. Skeuomorph versus flat design: user experience and age-related preferences. In: Marcus, A., Wang, W. (Eds.), Design, User Experience, and Usability: Designing Interactions, vol. 10919. Springer International Publishing, pp. 527–542. https://doi.org/10.1007/978-3-319-91803-7_40.

Belveal, R., 2013. Where have all the affordances gone? March 19. http://belveal.net/2013/03/19/where-have-all-the-affordances-gone/.

Benensohn, R., 2015. An open letter to "minimalist" sites. November 30. https://sitetuners.com/blog/an-open-letter-to-minimalist-sites/.

Blackwell, A.F., 2006. The reification of metaphor as a design tool. ACM Trans. Comput. Hum. Interact. 13 (4), 490–530. https://doi.org/10.1145/1188816.1188820.

Blaynee, J., Kreps, D., Kutar, M., Griffiths, M., 2016. Collaborative HCI and UX: longitudinal diary studies as a means of uncovering barriers to digital adoption. In: BCS Human Computer Interaction Conference 2016. https://doi.org/10.14236/ewic/HCI2016.72.

Bradley, S., 2013. Flat design done wrong. July 8. http://vanseodesign.com/web-design/flat-design-done-wrong/.

Burmistrov, I., Zlokazova, T., Izmalkova, A., Leonova, A., 2015. Flat design vs traditional design: comparative experimental study. In: Abascal, J., Barbosa, S., Fetter, M., Gross, T., Palanque, P., Winckler, M. (Eds.), Human-Computer Interaction – INTERACT 2015, vol. 9297. Springer International Publishing, pp. 106–114. https://doi.org/10.1007/978-3-319-22668-2_10.

China Internet Network Information Center, 2019. Statistical report on the development of Internet in China. http://www.cac.gov.cn/2019-08/30/c_1124939590.htm.

Cho, M., Kwon, S., Na, N., Suk, H.-J., Lee, K., 2015. The elders preference for skeuomorphism as app icon style, 899–904. https://doi.org/10.1145/2702613.2732887.

Creager, J.H., Gillan, D.J., 2016. Toward understanding the findability and discoverability of shading gradients in almost-flat design. Proc. Hum. Factors Ergon. Soc. Annu. Meet. 60 (1), 339–343. https://doi.org/10.1177/1541931213601077.

Gu, B., 2016. East Meets West on Flat Design: Convergence and Divergence in Chinese and American User Interface Design, vol. 17.

Haramundanis, K., 1996. Why Icons Cannot Stand Alone, vol. 8.

Hassenzahl, M., Burmester, M., Koller, F., 2003. AttrakDiff: ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In: Szwillus, G., Ziegler, J. (Eds.), Mensch & Computer 2003, vol. 57. Vieweg+Teubner Verlag, pp. 187–196. https://doi.org/10.1007/978-3-322-80058-9_19.

Hornor, T., 2015. Why the flat design trend is hurting usability. August 7. https://www.vandelaydesign.com/why-the-flat-design-trend-is-hurting-usability/.

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. Factors: J. Hum. Factors Ergon. Soc. 49 (3), 465–476. https://doi.org/10.1518/001872007X200102.

Kreps, D., Burmeister, O.K., Blaynee, J., 2016. Skeuomorphic reassurance: personhood and dementia. In: Kreps, D., Fletcher, G., Griffiths, M. (Eds.), Technology and Intimacy: Choice or Coercion, vol. 474. Springer International Publishing, pp. 61–71. https://doi.org/10.1007/978-3-319-44805-3_6.

Leung, R., McGrenere, J., Graf, P., 2011. Age-related differences in the initial usability of mobile device icons. Behav. Inf. Technol. 30 (5), 629–642. https://doi.org/10.1080/01449290903171308.

Linowski, J., 2013. Calling Bull$#!%: on flat design. May 1. http://wireframes.linowski.ca/2013/05/calling-bull-on-flat-design/.

McDougall, S.J.P., Reppa, I., 2008. Why do I like it? The relationships between icon characteristics, user performance and aesthetic appeal. In: Nd ANNUAL MEETING, 5.

McDougall, S.J.P., Curry, M.B., de Bruijn, 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 (3), 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 (4), 291–306. https://doi.org/10.1037//1076-898X.6.4.291.

Moran, K., 2015a. The characteristics of minimalism in Web design. July 12. https://www.nngroup.com/articles/characteristics-minimalism/.

Moran, K., 2015b. Flat design: its origins, its problems, and why flat 2.0 is better for users. September 27. https://www.nngroup.com/articles/flat-design/.

Moran, K., 2016. Young adults appreciate flat design more than their parents do. February 8. https://www.nngroup.com/articles/young-adults-flat-design/.

Moran, K., 2017. Flat UI elements attract less attention and cause uncertainty. September 3. https://www.nngroup.com/articles/flat-ui-less-attention-cause-uncertainty/.

Nielsen, J., 2012. Windows 8—disappointing usability for both novice and power users. November 19. https://www.nngroup.com/articles/windows-8-disappointing-usability/.

Page, T., 2014. Skeuomorphism or flat design: future directions in mobile device User Interface (UI) design education. Int. J. Mobile Learn. Organisat. 8 (2), 130. https://doi.org/10.1504/IJMLO.2014.062350.

Pan, Y., Stolterman, E., 2015. What if HCI became a fashion-driven discipline? Interactions 22 (6), 50–53. https://doi.org/10.1145/2828430.

Saffer, D., 2005. The Role of Metaphor in Interaction Design, vol. 29.

Schröder, S., Ziefle, M., 2008. Effects of icon concreteness and complexity on semantic transparency: younger vs. Older users. In: Miesenberger, K., Klaus, J., Zagler, W., Karshmer, A. (Eds.), Computers Helping People with Special Needs, vol. 5105. Springer, Berlin Heidelberg, pp. 90–97. https://doi.org/10.1007/978-3-540-70540-6_12.

Spiliotopoulos, K., Rigou, M., Sirmakessis, S., 2018. A comparative study of skeuomorphic and flat design from a UX perspective. Multimodal Technol. Interact. 2 (2), 31. https://doi.org/10.3390/mti2020031.

Svennerberg, G., 2012. Design for clickability. November 18. http://www.svennerberg.com/2012/12/design-for-clickability/.

Tao, J., Yafeng, N., Lei, Z., 2017. Are the warning icons more attentional? Appl. Ergon. 65, 51–60. https://doi.org/10.1016/j.apergo.2017.05.012.

Think with Google, 2016. How people discover, use, and stay engaged with apps. https://www.thinkwithgoogle.com/advertising-channels/apps/app-marketing-trends-mobile-landscape/.

Treder, M., 2013. 5 Dangers of flat design. June 28. https://www.uxpin.com/studio/blog/5-dangers-of-flat-design/.

Weinswig, D., 2016. Millennials go minimal: the decluttering lifestyle trend that is taking over. September 7. https://www.forbes.com/sites/deborahweinswig/2016/09/07/millennials-go-minimal-the-decluttering-lifestyle-trend-that-is-taking-over/#2fd413537550.

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 (2), 68–82. https://doi.org/10.1080/014492999119129.

Wu, L., Lei, T., Li, J., Li, B., 2015. Skeuomorphism and flat design: evaluating users' emotion experience in car navigation interface design. In: Marcus, A. (Ed.), Design, User Experience, and Usability: Design Discourse, vol. 9186. Springer International Publishing, pp. 567–575. https://doi.org/10.1007/978-3-319-20886-2_53.

Zhang, X., Wang, Q., Shi, Y., 2017. Contrastive analysis on emotional cognition of skeuomorphic and flat icon. In: Zhao, P., Ouyang, Y., Xu, M., Yang, L., Ouyang, Y. (Eds.), Advanced Graphic Communications and Media Technologies, vol. 417. Springer, Singapore, pp. 225–232. https://doi.org/10.1007/978-981-10-3530-2_28.

Zhou, J., Chourasia, A., Vanderheiden, G., 2017. Interface adaptation to novice older adults' mental models through concrete metaphors. Int. J. Hum. Comput. Interact. 33 (7), 592–606. https://doi.org/10.1080/10447318.2016.1265827.