

Experimental Investigation of Gender-Based Preferences for Skeuomorphic Flat Icons via Eye-Tracking: Evidence from 18-28-year-old Participants

Qingli Liu^{a,*}, Chuhan Xu^a, Haojing Guo^b

^a Jiangsu University, Zhenjiang 212013, China;

^b Department of Design, North Carolina State University, Raleigh, NC 27606, USA;

Abstract: To address the lack of objective physiological data and theoretical analysis on the impact of gender on preference for skeuomorphic and flat icons in existing research, the experiment recruited 15 male and 22 female participants aged 18–28 from Chinese universities. They conducted intra-group testing on 10 sets of functionally matched icons, which had been reviewed by professional designers to ensure consistency ($Kappa \geq 0.6$), and used SPSS 26.0 for data analysis. The results showed that in eye-tracking metrics, the main effect of icon style is significant. Skeuomorphic icons exhibit shorter initial fixation duration, longer total fixation duration, and greater total saccade amplitude. There was no significant interaction effect between gender and style. In terms of subjective preference, selection frequency of skeuomorphic icons was negatively correlated with gender ($r = -0.351$, $p = 0.033$), with males showing greater recognition of the attractiveness and intuitiveness of skeuomorphic designs, while females demonstrated stronger preference for the simplicity and modernity of flat designs. This difference correlated with variations in gender aesthetic orientation and functional focus weight, aligning with Norman's three-level logic of emotional design. The study provides empirical evidence for gender-sensitive icon design among young adult users in Chinese universities and expands the application pathways of emotional design theory in interface evaluation.

Keywords: Icon Design; Eye-tracking Experiment; Visual Preference; Affective Design, Gender Cognition

1. Introduction

1.1 Icon Style

Since digital human-computer interfaces have evolved so greatly, graphical symbols have become the primary means of conveying functional instructions and status feedback, and icons play a decisive role in information recognition, operation decision-making, and interaction efficiency[1]. More importantly, as interface-specific visual signs, application icons explicitly denote app functions and thus naturally support fast user navigation and selection[2]. In the high-frequency, information-dense mobile context, users must recognize icon meanings and complete operations within very limited time, which means that icons are no longer mere decoration: they directly affect interface usability, learning cost, and task efficiency[3]. Finally, while icon design must satisfy the basic requirements of comprehensibility and recognizability, it is also true that stylistic elements are legitimate and valuable design components in their own right[4].

Since icon styles can reasonably be divided into two main categories, namely skeuomorphic and flat design, reference may be discussed in[5]. With the development of internet and mobile application ecosystems, graphical user interfaces have become ubiquitous in usage scenarios, which have become two representative interface design paradigms[6]. Skeuomorphic icons emphasize faithfully reproducing real-world objects or contexts through details such as shapes, colors, textures, and surface finishes, characterized by rich details, realistic textures,

and pronounced light effects. For UI functional controls like knobs, buttons, sliders, and switches, these elements often draw visual inspiration from physical device prototypes[7], providing users with sensory cues akin to the real world to reduce comprehension barriers and enhance functional recognition. In contrast, flat design icons minimize texture and volumetric effects while retaining core contours and basic shapes, presenting information through clean lines and solid color blocks without shadows or perspective effects, pursuing more abstract and symbolic visual expressions.

It is worth noting that skeuomorphic and flat design strike different balances between realistic sensory hints and streamlined symbolic expression. Skeuomorphic icons can quickly activate associations rooted in users' past experience, whereas flat design helps cut down on visual interference and boosts the overall unity of dense, complex interfaces. Today's application interfaces generally demand both efficient recognition and strong visual appeal. For this reason, the choice between the two styles in practical design is not just a matter of aesthetic taste; it also relates to user groups, task scenarios, and how people process visual information[8]. Apart from basic functional usability, visual aesthetics play a key role in building user trust toward digital interfaces and sustaining long-term user engagement. Interface design features directly mold how users perceive and evaluate products[9]. A systematic analysis of how users perceive these two icon styles therefore delivers clear, testable guidance for icon design. It also improves the scientific validity and generalizability of decisions about interface styling [10].

1.2 Gender Cognition Differences

The combined influence of physiological structure and long-term experiential construction may lead to gender-specific differences in visual attention allocation, information processing strategies, and emotional cue extraction. Research in cognitive psychology and neuroscience consistently indicates that gender factors significantly impact individuals' visual perception and information processing patterns. Existing studies demonstrate that males typically exhibit relative advantages in spatial perception, directional localization, and spatial manipulation tasks[11], while females demonstrate greater sensitivity in color perception, detail recognition, and emotional cue detection[12]. Additionally, females may demonstrate a stronger tendency to identify stable patterns and natural object features under specific stimuli[13]. These gender-related visual preference differences have been validated in interface design applications such as web interfaces and smart wearables devices[14], providing empirical foundations for research on gender preferences in icon design.

In interface icon contexts, gender differences do not necessarily manifest as one-sided preferences for specific styles, but rather reflect variations in attention weighting toward different visual cues. As functional entry points, icons' recognition efficiency and interpretability directly impact task completion processes. Male users tend to prioritize functional usability and operational efficiency when selecting icons, while female users may focus more on visual ambiance, emotional affinity, and experience satisfaction[15]. Empirical studies further demonstrate that males exhibit significantly greater preference for flat icons compared to females, who generally favor minimalist and clean visual designs[16]. When icon styles show distinct differences in detail richness, realism cues, and abstraction levels, these attentional weightings may further influence users' initial attention patterns, comprehension pathways, and final preferences. This makes the selection differences between skeuomorphic and flat icons a research question worthy of empirical investigation.

Current research reveals three persistent shortcomings. Firstly, most studies focus on general preferences or single performance metrics while inadequately addressing evidence regarding the attention processes underlying these preferences, resulting in explanations that remain confined to superficial results. Secondly, gender differences are often treated as descriptive variables in samples rather than being integrated into core analytical frameworks to examine their interactive effects with icon styles. Thirdly, in young user demographics, both skeuomorphic and flat design styles coexist within real-world application ecosystems, where aesthetic judgments, functional comprehension, and emotional associations may occur simultaneously. However, existing research

rarely integrates these dimensions within a unified theoretical framework. Addressing these gaps, this study introduces a three-tiered model of emotional design, constructing coherent explanations across instinctive, behavioral, and reflective levels, while validating findings through eye-tracking data and subjective evaluations.

1.3 Three-level Model of Emotional Design

Donald A. Norman's Affective Design Model presents a clear, elegant framework for analyzing product user experience, namely the three hierarchical levels: the visceral level, the behavioral level, and the reflective level. The model also makes an explicit and well-supported claim: design that uses visual means such as patterns, colors, materials, and formal language shapes perceptual experiences, elicits emotional responses, facilitates meaning construction in human-computer interactions, and thus leads naturally to a multi-dimensional evaluation of the experience [17]. The visceral layer is fundamentally about instant emotional reactions to sensory input, namely visual, auditory, and tactile stimuli, so for icon design naturally corresponds to users' first impressions of visual properties such as hue, form, contrast, texture, and light effects. Therefore, responses at this layer typically manifest as early attention capture and rapid allocation of cognitive resources. In contrast, the behavioral layer is concerned with operational usability and task completion efficiency in actual use, which translates directly to easy functional recognition and simple operation in icon design. Skeuomorphic icons expertly use real-object metaphors and affordance hints to reduce cognitive barriers, making their functions immediately and intuitively clear to novice users [18]. Because flat icons eliminate decorative elements and visual clutter and therefore express symbols more clearly, users who form stable mental associations with them can process information more quickly and perceive interface presentations as more consistent [1,19]. The reflective layer incorporates advanced cognitive processes such as value evaluation, meaning connection, and emotional resonance, all of which are shaped by personal experience, cultural context, and aesthetic inclination. Because skeuomorphic icons closely resemble real-world objects, they are very likely to elicit nostalgia and a feeling of familiarity, in sharp contrast to flat icons, which are clearly linked to modernity, minimalism, and tech aesthetics, and thus better meet users' needs for neat visual layouts and up-to-date design sensibilities. More importantly, experiences at the reflective layer are ultimately grounded in the visceral and behavioral layers and are inevitably influenced by personal mental states, usage contexts, and cultural backgrounds[20].

Based on Norman's three-tiered model, this paper organizes the influence path of icon style differences from the visceral-level visual attraction to the behavioral-level functional utility, and then to the reflective-level emotional and cultural association. Accordingly, a corresponding analysis of subjective evaluation dimensions and objective eye-tracking attention metrics is conducted, enabling the interpretation of experimental results to advance from the descriptive level of 'what users like' to the explanatory level of 'why they like it and how preference is formed'.

1.4 Research Objectives and Hypotheses

This study investigates the differences in icon selection and cognitive processing among young users across two design styles—skeuomorphic and flat icons—while examining whether gender differences moderate how icon styles affect attention allocation, functional recognition, and emotional evaluation. Building on Norman's three-tiered Affective Design Model, we decompose icon user experience into three core dimensions: the visceral-level initial visual attraction, the behavioral-level functional understanding and intuitive usability, and the reflective-level aesthetic association and modernity perception. By combining objective eye-tracking metrics with subjective user evaluations, we establish a verifiable multi-dimensional evidence chain. Through correlation analysis between subjective ratings and eye-tracking data, the study explores how different icon styles function in the attention capture—comprehension facilitation—preference formation chain, and whether these roles exhibit gender-specific weight structures. The findings aim to provide more explanatory empirical evidence for interface icon style

selection in human-computer interface design.

From the present discussion, it is natural and appropriate to formulate the following research hypotheses.

H1: Since skeuomorphic icons are more likely to catch the user's initial attention and engage the user visually compared to flat icons, it follows that they have better attention-grabbing properties.

H2: Since icon style has a definite and important effect on functional recognition experience, it is natural and logical to say that skeuomorphic icons, which contain more rich real-world metaphorical cues, are better suited for intuitiveness evaluation and functional comprehension, whereas flat icons are better at reducing information processing load and improving interface clarity.

H3: Because flat icons are generally regarded as modern and minimalist in the reflection layer, whereas skeuomorphic icons have a clear tendency to evoke familiarity and nostalgia, the two types are accordingly evaluated quite differently with respect to appeal and emotion.

H4: The relationship between icon style and its effects on attention and evaluation is clearly moderated by gender: male users are more sensitive to functional cues and usability characteristics of skeuomorphic icons, hence they show a stronger preference for skeuomorphic designs in judgments about intuitiveness and visual appeal. In contrast, female users are more sensitive to simplicity, modernity, and the general experiential atmosphere, so they favor flat design styles in evaluations relating to simplicity and contemporary aesthetics.

H5: Since there is a clear, consistent relationship between subjective preference and eye-tracking attentional characteristics, it is natural to find that icon styles selected or rated higher are associated with greater attentional resource allocation, hence providing direct evidence for the attention-to-preference process.

2. Method

2.1 Experimental Design

The present study used a within-subjects design, with icon style (skeuomorphic/flat) and participant gender as the independent variables, and combined objective eye-tracking data with subjective rating scales to systematically assess differences in users' cognitive processing and preferences regarding interface icons. Eye-tracking was a non-intrusive method for recording visual exploration, and therefore it became a standard tool for measuring attention distribution and information processing in both physical and digital interfaces[21]. Corresponding to the three-tier Affective Design Model mentioned earlier, this study established four categories of eye-tracking metrics and two types of subjective indicators: First fixation duration (FFD) reflected the initial capture capacity of stimuli, indicating instant attraction and attention initiation at the visceral level[22]; Total fixation duration (TFD), represented cumulative gaze duration during individual trials or tasks, reflecting attention resource allocation and cognitive processing depth on target information, commonly used to measure comprehension load and processing intensity; First saccade direction(FSD) referred to the initial eye movement direction and target point after stimulus presentation, helping assess users' initial preference and attention transfer tendencies toward the two icon types; Total saccadic amplitude (TSA), representing cumulative saccadic angles or pixel distances within specific time intervals, demonstrates visual search scope and exploration intensity, providing process-level explanations for browsing strategy differences behind selection outcomes. The subjective data section included icon selection frequency and Likert scale scores, with the former directly reflecting preference orientation and the latter evaluating user experience from dimensions such as attractiveness, intuitiveness, simplicity and modernity, offering interpretable psychological mappings for objective eye-tracking metrics. Regarding experimental equipment and materials, this study employed an eye tracker to collect visual behavior data, supplemented by post-experiment subjective questionnaires and semi-structured interviews. The eye tracker captured participants' fixation and saccade patterns during screen-based experimental tasks, with data recording

and metric extraction performed by dedicated professional software. The subjective questionnaire utilized a 5-point Likert scale to evaluate four dimensions of user experience—attractiveness, intuitiveness, simplicity and modernity—for both skeuomorphic and flat icons. Post-experiment brief interviews were conducted to identify participants' primary decision-making rationale and key emotional experiences, providing interpretive context for quantitative experimental results. These tools and items collectively formed an evidence chain of process data (eye tracking)—outcome data (selection behavior)—interpretive data (ratings and interviews), supporting subsequent testing of icon style effects and gender differences.

2.2 Experimental Materials

We chose 10 representative icon sets, each containing one skeuomorphic icon and one flat icon, and systematically eliminated background interference by using a uniform #FFFFFF white background[23]. The skeuomorphic icons were designed in strict compliance with Nielsen's Usability Design Principles [24], whereas the flat icons followed Google Material Design guidelines. All icons in different sets served the same functional purpose, were all sized exactly 128 × 128 pixels, and had the same resolution of 72 dpi, as illustrated in Figure 1. Consequently, the experimental interface was designed with clear comparability in both functionality and usage context.



Figure 1. Sample of Skeuomorphic Icon and Flat Icon

Since the purpose of the study was to verify the effectiveness of skeuomorphism and flat design as experimental variables, it accordingly carried out a professional review by designers to evaluate the consistency of two icon sets along four clearly defined dimensions: line uniformity, color harmony, complexity balance, and metaphorical coherence[25], as described below.

We chose five professional designers who have at least three years of UI/UX design experience in order to eliminate random bias and properly satisfy the basic requirements of the Kappa test: two doing skeuomorphic design, two doing flat design, and one with cross-styled experience.

A 5-point scale was used to conduct cross-dimensional consistency evaluation for each set of icons:

Table 1. Kappa Consistency Analysis Scale

Dimension	1 Point (Poor)	3 Points (Moderate)	5 Points (Excellent)
Line Unity	Significant line width/curvature difference (>30%)	Partial inconsistency in line type (10%–20% difference)	Highly unified line style for all icons (<5% difference)
Color Harmony	Obvious hue/luminance jump	Individual icons with abrupt colors	Smooth color transition
Complexity Balance	Significant variation in element quantity (range≥3)	Most icons meet the complexity rule (range=1)	Strictly consistent number of elements for all icons (range=0)

Dimension	1 Point (Poor)	3 Points (Moderate)	5 Points (Excellent)
Metaphorical Consistency	Functional expression conflicts with industry conventions	Some icons require secondary cognitive interpretation	Symbolic metaphor in accordance with international standards

Kappa consistency analysis (Cohen's Kappa/Fleiss' Kappa) is a statistical measure of inter-rater reliability, with its core formula being:

$$K = \frac{o-e}{1-e} \tag{1}$$

Po: Observed Agreement

Pe: Theoretical probability of random agreement (Expected Agreement by Chance)

k: Final consistency coefficient (range: -1 to 1)

Table 2. Kappa Consistency Analysis Score Table

Evaluation Dimension	Skeuomorphism Group	Flat Group	Criterion
Line consistency (Kappa)	0.67	0.73	≥0.6
Color Harmony (Kappa)	0.62	0.68	≥0.6
Complexity balance (Kappa)	0.65	0.66	≥0.6
Metaphorical consistency (Kappa)	0.72	0.64	> 0.6
Inter-group differences (t-test p-value)	0.12	0.09	p>0.05 (no significant difference)

The Kappa analysis results (all Kappa values ≥ 0.6) demonstrated that the group icons maintained high consistency across their respective style dimensions, with no significant differences in factors such as complexity between different style groups (p > 0.05). This confirmed that the experimental materials successfully established visual contrast between skeuomorphic and flat styles while effectively controlling irrelevant variables such as functional consistency.

2.3 Participants

Previous research indicates that age significantly influences users' icon style preference patterns. Younger users tend to appreciate the modern aesthetic conveyed by flat icons, while older users rely more on the high recognizability provided by skeuomorphic icons[26]. Skeuomorphic icons demonstrate superior recognition accuracy and faster processing efficiency compared to flat icons because they can provide more sensory and visual cues; through realistic design, users can more easily identify icon types and their core meanings. Consequently, elderly individuals and children exhibit a stronger preference for skeuomorphic icons than younger demographics [18]. Conversely, younger users demonstrate higher familiarity with app icons, exhibit greater sensitivity to visual cues than older generations and children, and show stronger recognition of flat icons[27]. To ensure experimental validity, the study exclusively enrolled young participants to eliminate potential confounding effects of age on research outcomes.

A pre-test questionnaire was used to screen 15 males and 22 females aged between 18-28 years, totaling 37 participants. All subjects had been using smartphones for over 6 years without any difficulties, with normal visual acuity or corrected visual acuity. Individuals suffering from conditions such as color blindness or color vision deficiency were excluded.

2.4 Experimental Equipment and Procedures

The experiment utilized the Tobii Pro Lab eye tracker and its accompanying software. The specific procedures are as follows:

(1) Preparation and Calibration: Participants were seated at a distance of approximately 60 cm from the display screen. The eye tracker was calibrated using the 9-point calibration method of Tobii Pro Lab (with additional parameters), with a tracking error $\leq 0.5^\circ$.

(2) Practice session: Participants completed one set (containing two icons) of informal experimental tasks to familiarize themselves with the experimental procedure and response requirements.

The procedure of a single trial was as follows:

A '+' fixation point appeared in the center of the screen for 500 milliseconds.

Simultaneously, a set of icons (one representing skeuomorphism and one representing flat design) was displayed for 10 seconds.

After the icon was displayed, participants had to make their selection by pressing the left and right keys on the keyboard within a 10-second observation period. The icon remained visible until the end of the 10-second period after selection, and the trial was not terminated prematurely.

Throughout the entire process, the eye tracker continuously recorded data. The recorded parameters included initial fixation duration, total fixation duration, initial saccade direction, and total saccadic amplitude.

(3) Subjective questionnaire: Participants were asked to assess the 'attractiveness', 'intuitiveness', 'simplicity', and 'modernity' of different icon styles on a 1–5 Likert scale. Additionally, an optional open-ended question was included: 'Please describe the primary considerations when you select an icon style.'

2.5 Data Analysis

Statistical analysis was performed using IBM SPSS Statistics 26.0. Paired-samples t-tests were employed to compare differences in various eye-tracking metrics between skeuomorphic and flat icons; point-biserial correlation analysis was used to investigate the relationship between gender and skeuomorphic icon preference frequency; repeated-measures analysis of variance (ANOVA) was applied to analyze subjective rating data. The significance level was set at $\alpha = 0.05$ for all statistical analyses.

3. Results

3.1 Behavioral Choice and Subjective Rating Results

To examine the relationship between gender and preference for icon styles, the total number of times each participant selected skeuomorphic icons was summed to obtain the continuous variable 'skeuomorphic icon selection frequency,' which was then analyzed using point-biserial correlation with the binary gender variable (male = 1, female = 0). The results are presented in Table 3.

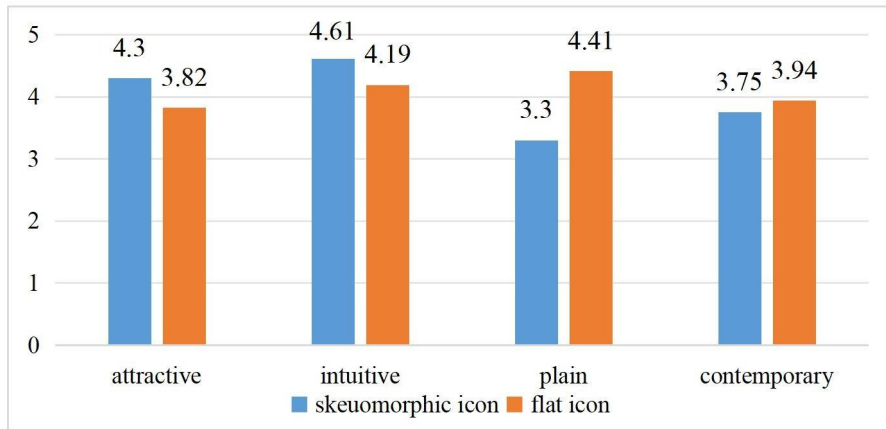
Table 3. Results of two-way correlation analysis (N=37)

Variable	Skeuomorphic Icon Selection Frequency
Gender (r_{pb})	-0.351
Sig. (2-tailed)	0.033

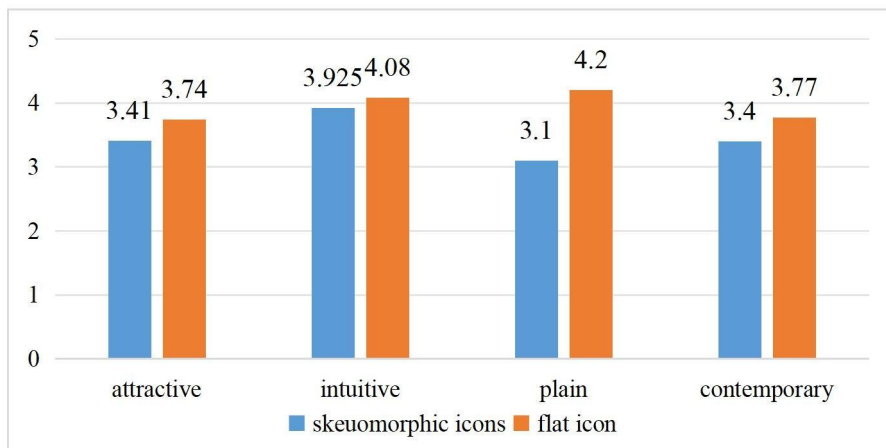
Note: r_{pb} = point-biserial correlation coefficient; male = 1, female = 0; N = 37.

The analysis revealed a statistically significant negative correlation between gender and the selection frequency of skeuomorphic icons ($r = -0.351$, $p = 0.033$, $N = 37$). According to Cohen's effect size criteria[28], $|r| = 0.351$ indicates a moderate effect ($0.3 \leq |r| < 0.5$).

To minimize the impact of individual icon design variations on overall conclusions, we conducted chi-square tests of independence separately on the selection outcomes of each icon group to examine the association between gender and selection outcomes at the group level.



(a)



(b)

Figure 2. Gender Differences in Subjective Ratings of Different Icon Styles. (a) Male Participants; (b) Female Participants.

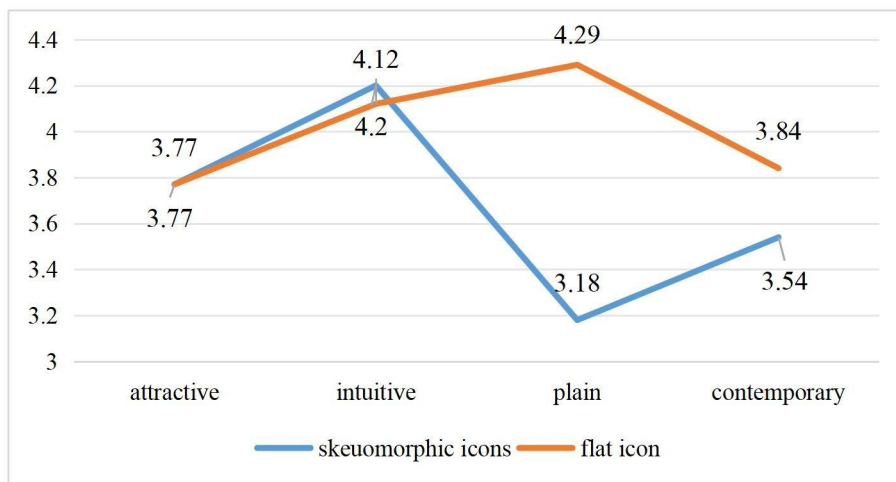


Figure 3. Overall Subjective Ratings of Skeuomorphic and Flat Icons Across Four Dimensions

In addition to behavioral choices, participants' subjective ratings on four dimensions — attractiveness, intuitiveness, simplicity, and modernity—were collected using the Likert scale for both skeuomorphic and flat

icons. The results are shown in Figures 2 and 3. Figure 2 illustrates the gender differences in subjective ratings for the two styles across these four dimensions, while Figure 3 summarizes the overall subjective rating distribution of both styles across all dimensions.

3.2 Eye-Tracking Index Results

To examine visual processing differences between male and female users under varying icon style conditions, this study conducted a statistical analysis of eye movement metrics and presented fixation distribution characteristics through heat maps. No critical eye movement indicators were missing during data preprocessing, with all 37 participants (15 males and 22 females) ultimately included in the analysis. To ensure comparability in repeated measurement analysis, eye movement metrics were analyzed using 'mean values across all style conditions' as statistical units. Icon style was treated as a within-subjects variable (skeuomorphism vs. flat design), while gender served as a between-subject variable (male vs. female).

3.2.1 Heatmap Results

From the analysis presented in Figure 4 of the spatial clustering characteristics of participants' gaze distribution, it is very clearly seen that the heat map shows a high concentration of fixations in the skeuomorphic icon area, whereas the flat icon area has more dispersed gaze clusters with no evident focal point.

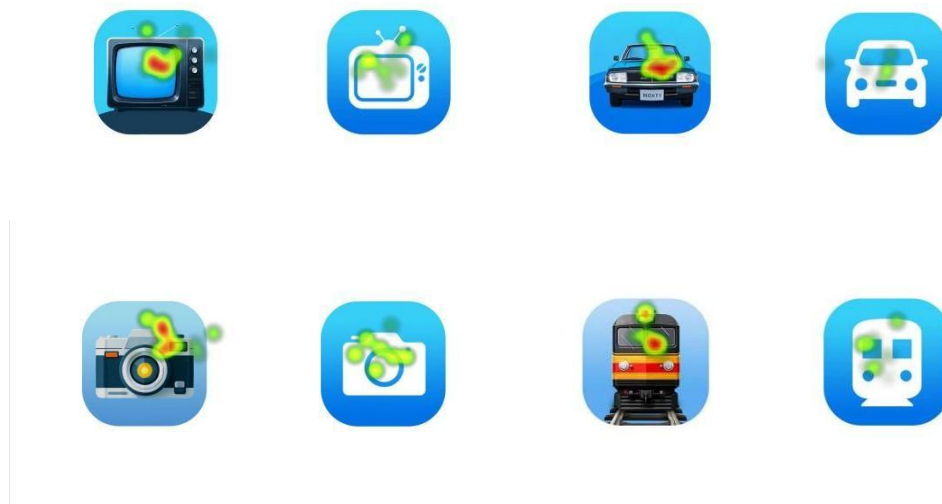


Figure 4. Heat Maps of the Eye-Tracking Experiment: Skeuomorphic Icon vs. Flat Icon

3.2.2 Descriptive Statistics of Key eye-tracking Parameters

Table 4 presents the descriptive statistics of key eye-tracking indicators for different genders under two style conditions. First Fixation Duration and Total Fixation Duration are measured in milliseconds (ms), Total Saccadic Amplitude is measured in angular degrees (°), and Saccade Count as the absolute number of saccades.

Table 4. Eye-Tracking Metrics of Different Genders Under Different Icon Style Conditions (M ± SD)

Metric	Male-Skeuomorphic	Male-Flat	Female-Skeuomorphic	Female-Flat
First Fixation Duration (ms)	242.55 ± 75.77	165.09 ± 41.15	228.20 ± 55.80	171.63 ± 55.37
Total Fixation Duration (ms)	2206.58 ± 927.74	1453.73 ± 492.40	1319.42 ± 609.66	1026.73 ± 385.46
Total Saccadic Amplitude (°)	70.54 ± 31.48	41.52 ± 17.76	60.16 ± 25.24	41.20 ± 21.36
Saccade Count (n)	6.20 ± 2.60	5.68 ± 2.19	4.61 ± 1.87	5.52 ± 2.30

3.2.3 Results of Two-Factor Repeated Measures ANOVA

Building upon the descriptive statistics, a 2 (Gender: Male/Female) \times 2 (Icon Style: Skeuomorphic/Flat) mixed-design repeated-measures ANOVA was conducted to analyze First Fixation Duration, Total Fixation Duration, Total Saccadic Amplitude, and Saccade Count. The sphericity assumption was automatically satisfied due to the two-level icon style factor, and partial eta-squared (η^2) was reported as the effect size.

Table 5. Summary of Two-Factor Repeated Measures ANOVA Results (N=37)

Metric	Effect	df	F	p	η^2
First Fixation Duration	Gender	1, 35	0.01	0.943	0.000
	Icon Style	1, 35	23.15	<0.001	0.398
	Gender \times Icon	1, 35	0.54	0.469	0.015
Total Fixation Duration	Gender	1, 35	3.3	0.078	0.086
	Icon Style	1, 35	10.29	0.003	0.227
	Gender \times Icon	1, 35	2.47	0.125	0.066
Total Saccadic Amplitude	Gender	1, 35	0.01	0.911	0.000
	Icon Style	1, 35	26.15	<0.001	0.428
	Gender \times Icon	1, 35	1.14	0.292	0.032
Saccade Count	Gender	1, 35	0.47	0.497	0.013
	Icon Style	1, 35	1.27	0.268	0.035
	Gender \times Icon	1, 35	2.3	0.139	0.062

The ANOVA results demonstrated that the main effect of Icon Style was statistically significant for three metrics: First Fixation Duration, Total Fixation Duration, and Total Saccadic Amplitude. Specifically, the main effect of Icon Style was significant for First Fixation Duration ($F(1,35)=23.15$, $p<0.001$, $\eta^2=0.398$), Total Fixation Duration ($F(1,35)=10.29$, $p=0.003$, $\eta^2=0.227$), and Total Saccadic Amplitude ($F(1,35)=26.15$, $p<0.001$, $\eta^2=0.428$). In contrast, the main effect of Icon Style for Saccade Count was not significant ($F(1,35)=1.27$, $p=0.268$, $\eta^2=0.035$).

Regarding gender factors, none of the four indicators showed a statistically significant main effect of Gender. The interaction effects of Gender \times Style for all indicators were not statistically significant ($p > 0.05$), indicating that the eye-tracking data did not reveal significant gender differences in the visual attention process. The prediction in the original Hypothesis H4 regarding 'gender moderating visual attention' was therefore not supported.

3.3 Qualitative Results of Post-Experiment Interviews

A thematic analysis was conducted on participants' responses to the open-ended question, 'Please describe the main factors you consider when choosing an icon style.' The results were highly consistent with Norman's three-level model:

At the visceral level (immediate visual experience): 8 participants mentioned the intuitive appeal of skeuomorphic icons derived from their texture, such as 'having a tactile feel' or 'being aesthetically pleasing/good-looking', corresponding to visceral-level processing.

From the behavioral level (function recognition and operational convenience) it is clearly and neatly established that 10 participants gave priority to function recognition and ease of use, and they explicitly recognized the intuitiveness and usability of skeuomorphic icons, such as 'you can tell the function at a glance'.

At the reflective level (aesthetic and value judgment), 19 participants judged flat design to be more modern

and minimalist from both aesthetic and value points of view, and similarly found skeuomorphic design to be cluttered, such as 'flat design is more modern/minimalist and clean, skeuomorphic design is too complex', hence this is a clear case of reflective-level processing.

4. Discussion

This study examined preference differences between skeuomorphic and flat icons through three evidence chains: behavioral choices, subjective evaluations, and eye-tracking processes. Behavioral data revealed a significant correlation between gender and selection frequency of skeuomorphic icons ($r = -0.351$, $p = 0.033$), indicating women generally chose skeuomorphic icons less frequently than men. Subjective ratings demonstrated distinct dimensional differentiation: flat icons scored higher in simplicity and modernity, while skeuomorphic icons tended to excel in appeal and intuitiveness among certain demographics. This 'dimensional differentiation' suggests that the two design styles do not represent a simple superiority hierarchy but rather highlight distinct advantages in experiential composition elements.

Eye movement data provides process-level supplementary evidence for the aforementioned preference differences. Two-way repeated measures ANOVA revealed that icon style had significant main effects on first fixation duration, total fixation duration, and total saccade amplitude, while the style main effect on saccade frequency was not statistically significant. Gender main effects were generally insignificant, and the gender-style interaction effect also failed to reach statistical significance. These results suggest that style differences exhibit relatively stable trends in influencing visual processing mechanisms, though such process variations did not demonstrate significant gender differentiation within the current sample conditions. Spatial distribution analysis using heat maps revealed a concentration pattern of fixation points in representationalized regions, consistent with the statistical findings.

Since the present study is in line with existing research frameworks, it is natural and appropriate to show its comparability with established conclusions about flat design's alignment with minimalist aesthetics and interface coherence, as well as skeuomorphism's reliance on tangible cues. It is argued in the text that previous research does not yet yield definite conclusions about the two design styles because their differences are inherently tied to task characteristics and evaluation metrics: flat designs with minimal visual noise are likely advantageous in visual search or efficiency-oriented tasks requiring rapid localization and interference suppression, whereas skeuomorphic details and material cues are more effective in preference selection and first impression judgments where interpretable cues and visual appeal are salient. Therefore, the selection+rating+eye-tracking combined paradigm used in this study is an ideal way to reveal the coexisting strengths of both design styles across multiple experiential dimensions.

In summary, the conclusions of this study can be summarized into three key points. At the behavioral choice level: Gender shows significant correlation with icon preference, with females demonstrating a stronger inclination toward flat design and males showing a relative preference for skeuomorphic design. At the subjective evaluation level: The two design styles exhibit dimension-specific differentiation, where flat design demonstrates superiority in simplicity and modernity, while skeuomorphic design excels in visual appeal and intuitiveness. At the visual attention level: Icon style significantly influences early visual processing, though the gender×style interaction effect remains statistically insignificant, indicating that gender differences do not manifest during the initial visual attention stage (instinctive/behavioral level) but emerge at later cognitive decision-making phases (reflective level). This finding directly does not support the original H4 hypothesis predicting gender-mediated regulation of visual attention. Preference formation constitutes a complex process, and final decisions do not directly reflect early visual processing patterns. The fact that behavioral preferences persist despite similar early visual processing mechanisms holds significant value. Preference differences occur at more advanced cognitive processing stages

(possibly corresponding to the reflective level in Norman's model), rather than being reflected through initial fixation patterns or total fixation duration at the instinctive/early behavioral level.

4.1 Theoretical and Practical Implications

Theoretically, this study integrates the three-level model of emotional design with eye-tracking evidence to establish a coherent explanatory framework for how icon styles influence user experience. The dimensional differentiation observed in subjective ratings, combined with the style main effects in eye-tracking metrics, demonstrates that icon styles affect not only aesthetic judgment but also engage multiple cognitive levels: attention initiation at the instinctive level, comprehension and processing engagement at the behavioral level, ultimately culminating in modern aesthetic associations at the reflective level. Notably, gender differences emerge more prominently in behavioral choices and evaluation tendencies, while showing no significant effects in interaction terms of key eye-tracking indicators. These findings provide a direction for future research to explore gender disparities through 'weight distribution in preference formation' rather than 'pure perceptual capability differences.' At the practical level, the research findings provide direct reference value for interface icon design and style decision-making among young student groups in China's universities. For scenarios emphasizing rapid understanding and functional prompts, skeuomorphism can be utilized to enhance intuitiveness; for high-information-density scenarios, flat design offers greater advantages. Icon design may adopt a layered strategy to balance efficiency and aesthetics.

4.2 Limitations of the Study

This study still faces limitations in boundary conditions. The sample size of only 37 participants (15 males) resulted in insufficient statistical power, leading to undetected eye movement interaction effects. The study population consisted of students from Jiangsu University with homogeneous educational backgrounds and a narrow age range of 18-28 years, making conclusions difficult to generalize to other age groups, cultural backgrounds, or technical experience levels. The stimulus materials comprised 10 pairs of skeuomorphic and flat icons. Although functional matching and consistency were controlled, factors such as icon familiarity, semantic distance, detail complexity, and color contrast may still influence attention metrics and subjective evaluations. Eye movement analysis primarily focused on aggregate indicators, with heat maps providing intuitive visualization but limited interpretive granularity. Without AOI-level modeling and stricter statistical controls, the robustness of inferences could be further enhanced. Additionally, gender was modeled as a binary variable, failing to account for more nuanced dimensions of gender identity and individual differences.

4.3 Future Research Directions

Future research can be expanded in three key areas. First, the lack of statistical power led to insignificant interaction effects in eye-tracking data, indicating the need for larger-scale studies with more balanced samples to validate or refute these findings. Research should incorporate more heterogeneous user groups across age and cultural backgrounds to test the stability and generalizability of conclusions through cross-sectional experiments. Second, systematic manipulation of stimulus materials—including icon complexity, semantic distance, familiarity, and depth effects—combined with comparative analyses of visual search, recognition, memory, and preference selection tasks, can better identify optimal conditions for different styles to demonstrate their advantages in efficiency and aesthetic dimensions. Third, analytical approaches should adopt AOI-level eye-tracking metrics and mixed-effects models, incorporating both participants and icon items as random effects in estimation. This methodology will enable rigorous testing of gender and style main effects and interactions, while integrating subjective ratings with eye-tracking metrics within a unified statistical framework to enhance conclusion interpretability and robustness.

5. Conclusions

This paper addresses the gap in the literature regarding the lack of objective physiological data and theoretical analysis on gender differences in skeuomorphic and flat icon preferences, by reporting an eye-tracking experiment with 37 participants aged 18-28. The results clearly and convincingly show that male users tended to select skeuomorphic icons more frequently than female users, as reflected by the significant negative correlation between gender and skeuomorphic icon selection frequency ($r = -0.351$, $p = 0.033$). Eye-tracking metrics clearly and convincingly showed that skeuomorphic icons have greater attention-grabbing power, with more focused gaze distributions, and the reasons for this preference were linked to gender-specific cognitive traits and social role orientations, as well as Norman's three-tier emotional design model. Therefore, the paper provides solid empirical support for gender-sensitive icon design and offers a meaningful extension of affective design theory to interface evaluation. The obvious limitations are the small sample size and single cultural context, so future work should address both by using more diverse samples and refined eye-tracking analysis methods.

Informed Consent Statement

Informed consent was obtained from all subjects involved in the study. Written informed consent has been obtained from the patient(s) to publish this paper.

Data Availability Statement

The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1]. Shen, Z., Chen, T., Wang, Y., Li, M., Chen, J., & Hu, Z. (2024). Skeuomorphic or flat? The effects of icon style on visual search and recognition performance. *Displays*, 84, 102813.
- [2]. Gittins, D. (1986). Icon-based human-computer interaction. *International Journal of Man-Machine Studies*, 24(6), 519–543.
- [3]. Yan, R. (2011). Icon design study in computer interface. *Procedia Engineering*, 15, 3134-3138.
- [4]. Chi, S. C. F. (2002). Factors affecting the design of computer icons. *International Journal of Industrial Ergonomics*, 29(4), 211-218.
- [5]. Gatsou, C., Politis, A., & Zevgolis, D. (2012). The importance of mobile interface icons on user interaction. *International Journal of Computer Science & Applications*, 9(3), 92-107.
- [6]. Kim, S., & Lee, S. (2020). Smash the dichotomy of skeuomorphism and flat design: designing an affordable interface to correspond with the human perceptuomotor process. *International Journal of Human-Computer Studies*, 141, 102435.
- [7]. Page, T. (2014). Skeuomorphism or flat design: future directions in mobile device user interface (ui) design education. *International Journal of Mobile Learning and Organisation*, 8(2), 130-142.
- [8]. Urbano, I. C. V. P., Guerreiro, J. P. V., & Nicolau, H. M. A. A. (2020). From skeuomorphism to flat design: age-related differences in performance and aesthetic perceptions. *Behaviour & Information Technology*, 41(3), 452–467.

- [9]. Hsu, C. C. (2012). Comparison of gender differences in young people's blog interface preferences and designs. *Displays*, 33(3), 119–128.
- [10]. Webb, J. M., Sorenson, P. F., & Lyons, N. P. (1989). An empirical approach to the evaluation of icons. *ACM SIGCHI Bulletin*, 21(1), 87-90.
- [11]. Munion, A. K. , Stefanucci, J. K. , Rovira, E. , Squire, P. , & Hendricks, M. (2019). Gender differences in spatial navigation: characterizing wayfinding behaviors. *Psychonomic Bulletin & Review*, 26(6), 1933-1940.
- [12]. Liu, Y., Ding, Y., & Yang, C. (2025). Seek for harmony or dominance? Gender differences in consumer preference for color contrast. *Psychology & Marketing*, 42(7), 1879-1891.
- [13]. Wang, R. W. Y. , Ke, T. M. , Chuang, S. W. , & Liu, I. N. . (2020). Sex differences in high-level appreciation of automobile design-evoked gamma broadband synchronization. *Scientific Reports*, 10(1), 9797.
- [14]. Wang, J., & Hsu, Y. (2020). Gender difference in preference for Apple Watch dial interface. In S. Yamamoto & H. Mori (Eds.), *Human interface and the management of information. Designing information* (pp. 170–182). Springer.
- [15]. Canning, C. (2012). Gender, design and marketing: How gender drives our perception of design and marketing. *Journal of Marketing Management*, 28(13-14), 1642-1644.
- [16]. Zhang, X., & Wang, H. (2024). User preferences for icon design styles and their associations with personality and demographic. In C. Stephanidis, M. Antona, S. Ntoa, & G. Salvendy (Eds.), *HCI International 2024 Posters* (pp. 235–245). Springer.
- [17]. Norman, D. A. (2004). *Emotional design: why we love (or hate) everyday things*. Basic Books.
- [18]. Zhang, X., Wang, Q., & Shi, Y. (2017). Contrastive analysis on emotional cognition of skeuomorphic and flat icon. In *Advanced Graphic Communications and Media Technologies* (pp. 225–232). Springer.
- [19]. Mu, D., Huang, Y., Wang, Y., Yang, J., Li, J., & Kang, Z. (2022). Neumorphic or flat? Impact of icon depth on user attention and visual search efficiency. *International Journal of Industrial Ergonomics*, 89, 103281.
- [20]. Wu, L., & Cui, X. (2021). Design method of subway station wall decoration based on emotional design theory—Take the Exhibition Center Station of Suzhou Line 7 as an example. In *Advances in Industrial Design* (pp. 909–914). Springer.
- [21]. Guo, R., Kim, N., & Lee, J. (2024). Empirical insights into eye-tracking for design evaluation: Applications in visual communication and new media design. *Behavioral Sciences* (2076-328X), 14(12).
- [22]. Tangmanee, C. (2016). Fixation and recall of Youtube ad banners: An eye-tracking study. *International Journal of Electronic Commerce Studies*, 7, 49–76.
- [23]. Yu, N., & Ouyang, Z. (2024). Effects of background colour, polarity, and saturation on digital icon status recognition and visual search performance. *Ergonomics*, 67(3), 433-445.
- [24]. Nielsen, J. (1994). *Usability engineering*. Morgan Kaufmann.
- [25]. Collaud, R., Reppa, I., Défayes L., McDougall S., Henchoz, N., & Sonderegger, A. (2022). Design standards for icons: the independent role of aesthetics, visual complexity and concreteness in icon design and icon understanding. *Displays*, 74(c), 102290.
- [26]. Chen, X., Xue, C., Wang, H., & Zhang, Q. (2017). A design method of icon based on semantic research of universal symbols. In *ISPE Inc. International Conference on Transdisciplinary Engineering* (pp.1066-1073).
- [27]. Panchoojit, L., & Hongwarittorn, N. (2019). Age differences in menu item selection for smartphone: The effects of icon background colors and icon symbols. In *Proceedings of the 5th International ACM In-Cooperation HCI and UX Conference* (pp. 55–64). Association for Computing Machinery.
- [28]. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Lawrence Erlbaum Associates.