The effects of skeuomorphic interface design on the usability of AR mobile apps

Linxi Wang
Media Technology and Interaction Design
KTH Royal Institute of Technology
linxi@kth.se

Huawei Jiang
Media Technology and Interaction Design
KTH Royal Institute of Technology
huaweij@kth.se

ABSTRACT
As all the user experience are generated around the user interface, the usability resulting in its design is considered to be a key factor that influences the effectiveness of the user interface and the interactive system as well. Skeuomorphism was widely used in interface design then gradually replaced by flat design and later material design. In recent years, the rise of augmented reality and virtual reality is bringing skeuomorphism back to the interface design thinking. In this research, we want to answer the question that whether the skeuomorphic design has positive effects on the usability of the interface for AR mobile applications by conducting experiments based on prototypes of skeuomorphic design and flat design. The tested hypothesis was validated with the results of descriptive statistical analysis, a one-way ANOVA test, and interviews. Our goal is to draw a conclusion from the usability test for a better design decision.

AUTHOR KEYWORDS
Augmented reality; Skeuomorphic interface design; Flat interface design; Usability test

INTRODUCTION
The user interface (UI), in the industrial design field of human-computer interaction, is the space where interactions between humans and machines occur. As all the user experience are generated around the user interface, the usability resulting in its design is considered to be a key factor that influences the effectiveness of the user interface and the interactive system as well. A good user interface design facilitates accomplishing the interaction at hand without drawing unnecessary attention to itself. Skeuomorphism was widely used in interface design at the beginning stage of the interface design for the handheld device, we widely simulated interaction from the physical world. In the past decade, handheld mobile devices have become increasingly prevalent, over the years, researcher and designers are exploring and utilizing different kinds of interface design. Trends are changing but with the same aim - to support its usability, enhancing the interactions and improving the aesthetic appeal of the design [1]. Then it was gradually replaced by flat design and later material design. Skeuomorphism referred to a design principle inspired from the physical world. And at the beginning of the interface design for the handheld device, we simulated interaction from the physical world widely. Before the last redesign of iOS in 2013, Apple adapted the skeuomorphic interface design. Then jumped into later-trendy flat design embracing the limitation of 2D interface without gradients, textures, shadows and shapes, which became quite popular for year until Google's release of material design to overcome the problem that in flat design clickable elements get usually confused with non-clickable icons and texts by introducing some skeuomorphism in the most simplified design. Skeuomorphism is always augmented with some non-skeuomorphic design of the interface, we keep what works and get rid of what doesn't. And in recent years, the rise of augmented reality and virtual reality is bringing skeuomorphism back to the interface design thinking. In this report, we are going to explore the effects of skeuomorphic interface design on the usability of AR mobile apps.

THEORETICAL FRAMEWORK
The usability resulting from the design of a user interface (UI) is a key factor which influences the effectiveness. Skeuomorphism comes from the notion of the skeuomorph, which is "a derivative object that retains ornamental design cues from structures that were necessary for the original"[2]. It describes an object or feature that imitates the design of a similar artifact made from another material[3]. Our research is built on several studies that compare the skeuomorphic and flat design from a user experience perspective. The theoretical framework is mainly based on the one in the paper "A Comparative Study of Skeuomorphic and Flat Design from a UX Perspective"[3]. And to make the framework more concrete, we will refer to other literature and studies that are related to the impact of skeuomorphic design on usability, such as "The Elders Preference for Skeuomorphism as App Icon Style"[4], and "Skeu the Evolution: Skeuomorphs, Style, and the Material of Tangible Interactions"[5]. The aforementioned papers are researching the design of skeuomorphism but from a different perspective and the design example used in the studies are employed for a

different situation; however, some of the results are contradictory. What we do in this study is to research based on the existing researches theoretical framework about whether the skeuomorphic design can affect the user's performance or not.

RESEARCH QUESTION

The research question of this project is whether the skeuomorphic design has positive effects on the usability of the interface for AR mobile applications. To answer this question, first of all, we need to set the definition for the terms of skeuomorphic design and usability in this research project. As mentioned before, skeuomorphic design is designing an element of a graphical user interface which mimics a physical object. And when it comes to a user interface, skeuomorphism is generally aimed at creating a three-dimensional effect on a flat surface. In our project, the skeuomorphic design in augmented reality mobile app mainly reflects on the visual interface of the AR view. And the usability measured in this project includes three components which are learnability, error, and satisfaction.

With the re-definition above, we came up with the testable hypothesis to start further investigation. The primary hypothesis is that users learn better in the interface of skeuomorphic design in augmented reality mobile app. Another hypothesis is that users prefer skeuomorphic interface design in augmented reality mobile app than the flat interface design. We falsified the hypothesis by running a series of experiments through the usability test, by which we wanted to measure the effects of skeuomorphic interface design from the aspects of user satisfaction, learnability, and errors. Our goal is to draw a conclusion from the usability test for a better design decision.

METHODOLOGY

In this research, we want to explore the effects of skeuomorphic interface design on the usability of AR mobile apps compared with flat design. There are five quality components of Usability: Learnability, Efficiency, Errors, Memorability, and Satisfaction.[6] In our research, we measure user satisfaction, learnability, and errors which are more important in our case, since there is a forced tradeoff that it is almost impossible to optimize all usability measures at the same time. The project employed both empirical method and analytical method. Two Low-fidelity prototypes are used in this project, which is defined to represent skeuomorphic design and flat design based on our own experience and knowledge. The usability test is conducted by analytical methods. For the learnability, time of the first attempt on the task is collected. For errors, the number of error is collected. For satisfaction, as it is not clear to measure the satisfaction of users quantitatively, a brief interview after the experiment is used to collect the data for measuring satisfaction, which is a common usability method to test satisfaction.

Participants

Based on the research of Nielsen Norman Group, it should test at least 20 users in quantitative test and at least 5 users in qualitative test in a usability test to ensure the quality of the experiment[7]. In our case, due to the limitations of time and resources, we chose 10 participants in our test. There are numerous researches to analyze the necessary number of participants that allows for statistically significant results, which the range is between 8 to 25 participants. So 10 participants is also a reliable test population[8]. In our project, the ideal situation is that participants should be from all the ages. But, due to the limitations of resources, the ages of our participants are from 21 to 28. There are 6 males and 4 females among 10 participants which can count as a balanced sample.

Low-Fidelity Prototypes

In our project, we used two low-fidelity prototypes to run the test, since it requires less preparation compared with high-fidelity prototypes [9] and we can have more time to focus on experiment and analysis of this project. These two prototypes are designed and created based on an AR mobile application - AR dragon which is a popular application related to AR in IOS platform. This application is an AR pet simulator that users can interact with. In this step, two prototypes are defined as skeuomorphic design and flat design based on our understanding of design principles.

Figure 1. Skeuomorphic design and flat design

Experiment Design

In our usability test, we used between experiment design. 10 participants are divided into two groups, which means that 5 participants tested the skeuomorphic prototype, while others tested the flat one. So each participant only tests one prototype, which can eliminate the mutual effects between these two prototypes. The order of the test is totally random. Each participant conducted two tasks which are same in two prototypes. Task 1 is to feed the AR pet, while Task 2 is to play with the AR pet. The data of finishing time and errors is collected manually. After testing, we ask a closed question in Likert scale. After this question, we show them the prototype of another design and have a walkthrough of the tasks as they did in the experiment, then asks them their preference and reason.

Experiment Process

a. Moderator introduces the project and process to participants.
b. Moderator asks basic information about participant (age, gender, previous experience for AR applications). Meanwhile, recorder record the information.
c. Participant is trained to use AR applications and is required to use the prototype as a real AR application during the testing
d. Participant tests the tasks. Recorder record the finishing time and errors.
e. Moderator asks a closed question in Likert scale (Q: do you like the design in this prototype? can you give a score for it, from 1 to 7, 7 is the best?).
f. Moderator shows another prototype and lets participant walkthrough of the tasks as the did in the experiment.
g. Participant is asked his preference between the two designs and reason.

RESULTS & ANALYSIS

Data Collection

During the whole process, there will be a moderator who gives a brief introduction and instructions to the participants, and a recorder who takes down all the details and results. To avoid adding confounding variables to the experiment, the moderator and the recorder is the same person for all the participants. The collected data in each experiment process includes brief information (age, gender, previous experience) about the participant (anonymous), the time for finishing each task, and the error made in each task.

To collect the quantitative data from the experiment, we employed the Touchstone to exploring alternative designs of controlled laboratory experiments, which facilitates the process of creating experiments. We generated a CSV file to collect the data in the experiment. The CSV file records participant ID, block, trail, DesignVariable, FinishTime, Error, and Succeed or not. And with the CSV file, we can start to work on statistical analysis. To collect the qualitative data, we conducted a short interview in end of the experiment to collect their opinion about the prototypes of skeuomorphic design and flat design to dig out more "stories" about the interaction for generating the insight.

There will be a detailed discussion about it in the following.

Data Analysis - Learnability and Errors

We used R Studio for statistics to have an understanding of the mathematical characteristics of data and estimate the probability that our hypotheses are correct. The first step is to summarize the data by descriptive statistics methods to see whether the hypothesis can stand on the data sample collected in the experiment. Then we use the one-way ANOVA test to use the data sample to provide a description of the hypothesis on the whole population can be trusted and testing the effect of skeuomorphic design on the "learning time".

Descriptive statistics - summarize the data

The dataset contains the information of 10 people who took participation in the experiment on one of the design style. The aim of the study was to see which design style is better for finishing time (learnability), so the independent variable (group) is the design style (DeisgnVariable). We calculated means and standard deviations for FinishTime by DesignVariable to see whether the group means and standard deviations look similar or very different. Table 1 presents the data calculated in R.

Table 1. Means and Standard Deviations for FinishTime

Design Variable	Mean	Standard Deviations
Skeuomorphic	7628	3010.381
Flat	11949	4110.059

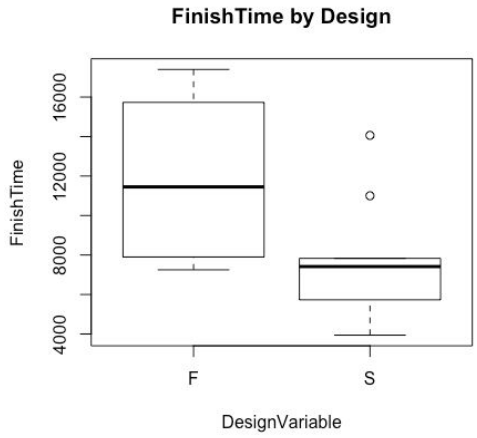

Figure 2.box plot of FinishTime by Design Variables

And on the right, a box plot is produced to depict the descriptive statistics of FinishTime by DesignVariable. From both Table1 and the boxplot, we can find that skeuomorphic seems better than "Flat" for the learnability of a user interface as the mean FinishTime is shorter. And there is also a noticeable difference in standard deviations.

One-way (between-groups) ANOVA in R

Rephrasing our hypotheses, we want to test the following:
Hypothesis: Users spend less time to finish the same task in skeuomorphic design interface than in flat design interface.
Null hypothesis: Users spend more time to finish the same task in skeuomorphic design interface than in flat design interface.
With the one-way ANOVA test, we want to see if there is a difference regarding the FinishTime measure between skeuomorphic interface design and flat design.
First, we create a summary table frame that contains the mean of FinishTime per participant x DeisgnVariable. Following are the results in R.

```
> data <- summarySE(UserTest2,measurevar= "FinishTime", groupvars=c("Participant","DesignVariable"))
> data
   Participant DesignVariable N FinishTime         sd   se        ci
1            0              S 2       7055  205.06097  145  1842.400
2            1              S 2       9195 6880.14898 4865 61815.686
3            2              S 2       7640   14.14214   10   127.062
4            3              S 2       9415 2241.52850 1585 20139.335
5            4              S 2       4835 1265.72114  895 11372.053
6            5              F 2       7330  113.13708   80  1016.496
7            6              F 2      14165 4475.98592 3165 40215.138
8            7              F 2      11385 4928.53426 3485 44281.124
9            8              F 2      12215 4970.96067 3515 44662.310
10           9              F 2      14650 3889.08730 2750 34942.063
```

Figure 3. summarize for ANOVA test

Second, we run the one-way ANOVA test in R, and the results are in Table 2. After getting the results of the one-way ANOVA test, we calculate the effect size which is a descriptive statistic that does not make any statement about whether the apparent relationship in the sample reflects a true relationship in the population.

From the results, we can observe a significant effect ($p = 0.0233 < 0.05$), which is the evidence for rejecting the null hypothesis. And we calculate the effect size which is a descriptive statistic that does not make any statement about whether the apparent relationship in the sample reflects a true relationship in the population. After this, we run a Tukey HSD post-hoc test to compares all possible pairs of means, in Table 3.

Table 2. Results from one-way ANOVA test.

	Df	Sum	Mean	F	Value Pr(>F)
DV	1	46677602	46677602	7.8	0.0233*
Residuals	8	47708250	5963531		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 3. Results from Tukey HSD post-hoc test

DV	diff	lwr	upr	p adj
S-F[1]	-4321	-7882.573	-759.4272	0.0232799

The ANOVA test revealed a significant effect of DesignVariable on FinishTime ($F(1,18) = 7.827$, $p = 0.023 < 0.05$, eta squared=0.49).
Hence, we reject the null hypothesis. The results indicate that the effect of DesignVariable on FinishTime is significant, which can support the results we got from the descriptive statistics. And with all these results, we validate the hypothesis that users spend less time to finish the same task in skeuomorphic design interface than in flat design interface. With this fact, we can prove that users spend less time to finish the same task in skeuomorphic design interface than in flat design interface from the perspective of learnability, which is the primary hypothesis from our research question.

User satisfaction

There is a Likert-type scale question after the experiment. Then we conducted a short oral interview of the participants' opinion toward the design style of the user interface. The process goes like this: we showed them the prototype of another design and had a walkthrough of the task as they did in the experiment, then asked them which kind of design do they prefer and the reason for preference. The results we got are in the following.

Score from Likert-type scale question

The average score of skeuomorphic design is 6.2 out of 7,

[1] S-F means Skeuomorphic-Flat.

and the average score of flat design is 4.2 out of 7. From the results of the satisfaction score, we can see that a noticeable difference between skeuomorphic design and flat design. Another fact is that all the answers for the question “which kind of design do you prefer? ” is skeuomorphic design for all the groups in the whole experiment. The reasons include “skeuomorphic design is prettier”, “it is easier to understand and more intuitive”, “it fits the AR view better”, “it is colorful and vivid”. And these reasons are mainly from two perspectives: aesthetics and perceived usability.

With what we got in this part, we can give a positive answer to our research question - whether the skeuomorphic design has positive effects on the usability of the interface for AR mobile applications.

DISCUSSION

There are still many limitations of the experiment process. The prototypes we used are based on a game-related app, which have the feature of adpatting skeuomorphic design. In the future, we should include other types of app to the experiment to get the results more convincing. Meanwhile, the low-fidelity prototype might influence the results. Also, different kinds of flat design may result in a different outcome. The design of tasks in the experiment can also influence the finishing time. And the way of timing is not accurate enough.

But anyway, with the results of the experiment, we can have the insight that the skeuomorphic design has positive effects on the usability of the interface for AR mobile applications. Nowadays, the rise of augmented reality is bringing skeuomorphism back to the interface design thinking. If the skeuomorphic design can help users perform better in AR mobile applications, designers could try to use more content related to skeuomorphism on AR mobile applications to improve the user experience.

REFERENCES

1. [1] Norman, D. A. (2002). "Emotion & Design: Attractive things work better". Interactions Magazine, ix (4). pp. 36–42. Retrieved 20 April 2014.
2. [2] Basalla, G. The Evolution of Technology; Cambridge University Press: Cambridge, UK, 1988.
3. [3] Spiliotopoulos, K.; Rigou, M.; Sirmakessis, S. A Comparative Study of Skeuomorphic and Flat Design from a UX Perspective. Multimodal Technologies Interact. 2018, 2, 31.
4. [4] Cho, M., Kwon, S., Na, N., Suk, H.-J., & Lee, K., 2015. The Elders Preference for Skeuomorphism as App Icon Style. In the Proceedings of the 33rd
5. [5] Gross, S.; Bardzell, J.; Bardzell, S. Skeu the evolution: Skeuomorphs, style, and the material of tangible interactions. In Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction (TEI 2014), Munich, Germany, 16–19 February 2014.
6. [6] Jokob Nielsen, 2012. Usability 101 : Introduction to Usability https://www.nngroup.com/articles/usability-101-introduction-to-usability/
7. [7] Jokob Nielsen, 2012. How Many Test Users in a Usability Study https://www.nngroup.com/articles/how-many-test-users/
8. [8] Macefield, R. How to specify the participant group size for usability studies: A practitioner’s guide. J. Usability Stud. 2009, 5, 34–45.
9. [9] Jokob Nielsen, 2016. UX Prototypes: Low Fidelity vs. High Fidelity https://www.nngroup.com/articles/ux-prototype-hi-lo-fidelity/