
Replicating and Extending Research on Relations between Visual Aesthetics and Usability

Noam Tractinsky

Ben-Gurion University of the Negev

Beer Sheva

Israel

noamt@bgu.ac.il

Abstract

This paper describes a replication and extension of a study that found strong positive correlation between evaluation of a product's beauty and pre-use perceptions of its usability. The original study was conducted in Japan; its replication and extension took place in Israel. The extension involved mainly methodological improvements to the original study, which demonstrated the robustness of the original study's findings.

Author Keywords

Replication; Visual Aesthetics; Perceived usability, Cross-culture, Method bias, HCI.

ACM Classification Keywords

H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces

Introduction and Motivation

At CHI '95, I attended a session in which Masaaki Kurosu presented a short paper. The paper described an experiment, designed to find whether people's perceptions of usability (operationalized as ease-of-use) correlate with established user interface design guidelines (Kurosu and Kashimura, 1995). Kurosu and

Presented at RepliCHI2013. Copyright © 2013 for the individual papers by the papers' authors. Copying permitted only for private and academic purposes. This volume is published and copyrighted by its editors.

Kashimura (K&K) projected 26 different designs of ATM interfaces to groups of participants seated in a classroom. During the projection of each slide the participants rated the design in terms of its perceived usability and beauty. Evaluations of usability were then correlated with the degree to which the designs followed usability guidelines. One of the study's results, however, pertained to the relation between the participants evaluations of usability and beauty (See Figure 1). To the best of my knowledge, that study was the first in the HCI literature to provide empirical evidence regarding the relation between these two system aspects. Surprisingly, the data indicated that people's perceptions of system's aesthetics are strongly and positively correlated ($r=0.59$) with their perceptions of the system's usability.

I was surprised by K&K's findings, and thought that their study should be replicated for several reasons. First, their results ran contrary to the prevailing thought in the field of HCI. At that time beauty (or visual aesthetics) was a marginal factor in HCI research and practice. It was usually ignored; rare acknowledgments of aesthetic design were immediately followed with caveats against overemphasizing it or with a message belittling its role relative to more utilitarian aspects and objectives of interactive systems.

Second, I was willing to accept that K&K's findings may hold in the particular locale of their study – Japan – a country with a long and glorious aesthetic tradition. However, I was skeptic about the generalizability of these findings to other places. More specifically, I found it unreasonable that similar correlations would be found in my own country – Israel – which is known more for

its people's orientation to act rather than for its aesthetic tradition.

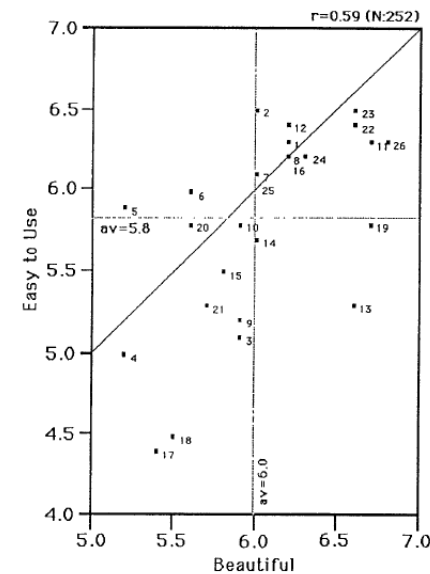


Fig.1 Correlation between two kinds of judgements for 26 layout samples.

Figure 1. Relationships between beauty and perceived usability as reported in Kurosu and Kashimura's (1995) study.

Finally, the method used in the original study was not flawless. In particular, there was a concern that the observed correlations between beauty and perceived usability were influenced, at least partially, by common method bias.

Thus, I embarked on a research project whose main objective was to demonstrate that K&K's findings were either wrong (as implied by main stream HCI

literature), or at best qualified by cultural factors (as implied by my own experience). The project, which is described below, included replication and extension of Kurosu and Kashimura's research. Its outcomes were published at CHI '97 (Tractinsky, 1997), and are summarized below.

Replication and Extension

My research included three studies: A replication of the original study and two extensions. All three studies used the same independent and dependent variables as the original study. The stimuli (designs of ATM machines) were basically the same as those of the original study, but had to be adapted to the locale of the replication studies. Whereas the first study replicated the original study's procedure, the next two studies extended it by employing increasingly more rigorous methods to examine the relationships between visual aesthetics and perceived usability.

Study 1 - Replication

Study 1 was an exact replication of K&K's method and stimuli with the exception that the Japanese stimuli had to be adapted to running the experiment in Israel. Most of the adaptation included the translation of the labels of certain controls of the ATM machine (e.g., the Confirm, Cancel, and Correction buttons). This part was quite simple, but there were two types of challenges. First, the original materials had to be reconstructed because of incompatible hardware and software. Second, while literal translation of the basic controls was straightforward, other parts of the interface were unique to Japan and were unfamiliar to Israeli users. For example, the original designs contained a large element depicting a feminine figure. This figure was

unique to Japanese ATMs. Israeli ATMs contained no similar element and it was feared that its inclusion would be met with skepticism (or worse). Thus, to prevent negative reactions on the one hand and to preserve the overall design layout on the other hand, the figure was replaced with a visual element of the same size, but which displayed an hour glass (see Figure 2, taken from Tractinsky, 1997).

Following the reconstruction of the stimuli the study followed the same procedure used in the original study.

Study 2 –Methodological Improvement I

Study 2 tested whether the results from the original study and its replication in Study 1 resulted from a method bias due to the fact that responses to the aesthetic and to the usability items were collected at the same time while the participants viewed the same design. That method carried the risk that the proximity of the measures would artificially inflate the correlation between them. To alleviate part of the concern, the study's procedure was modified. The 26 designs were displayed in two separated rounds. The order of presentation of the designs was randomized within each round. The order of evaluating beauty and ease of use was counterbalanced between two groups of participants.

Study 3 –Methodological Improvement II

In the original study and in the first two replication studies, the designs were presented to large groups of participants on a common screen, using a slide projector. In Study 3 the designs were presented on a computer screen by a program that also collected the

participants' responses. The use of computerized program allowed to further reduce potential biases by presenting the designs and the items measuring beauty and usability in a completely randomized order. The

differences between the three replicating studies are presented in Table 1.

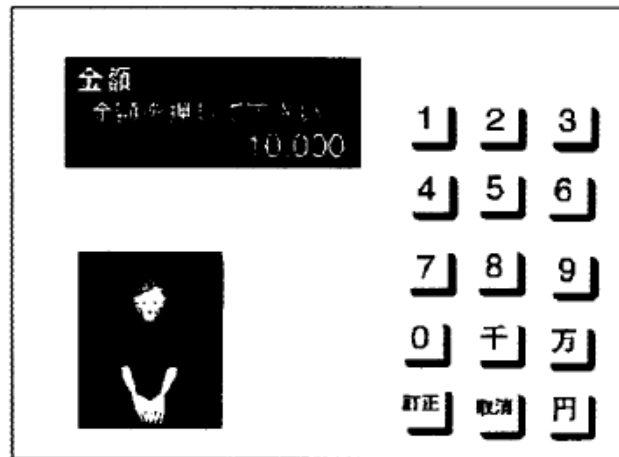


Figure 1(a). An original Japanese interface, rated high on apparent usability and aesthetics.



Figure 1(b). The equivalent Israeli interface, rated high on apparent usability and aesthetics.

Figure 2. Comparison of an original design and its counterpart in the replicating study

Study 1: Replication N = 104	Study 2: Improved method (1) N = 81	Study 3: Improved method (2) N = 108
Same procedure as original study. Designs adapted to fit local language/culture.	Items measuring beauty and usability were separated and their order of presentation was counter-balanced.	(a) Stimuli (designs) and measuring items were completely randomized. (b) Participants were seated individually in front of a computer.

Table 1. Summary of differences between studies of the replicating research.

Results

The three replicating studies yielded results similar to the original study in terms of the correlations between perceived (apparent) usability and six of the seven design guidelines, and most importantly, between perceived usability and evaluations of the designs' aesthetics. As can be seen in Figure 3, the basic findings remained unaffected by methodological improvements. If anything, the correlations between perceived usability and beauty were even higher in the replicating studies, demonstrating the robustness of the original findings.

Conclusion

The consistent results across cultures and following methodological improvements lent credibility to the findings of the original study. The original study and its replication opened up a new and lively research area in HCI regarding the role of visual aesthetics in HCI, and regarding its antecedents and consequences.

Variable	Correlations with Apparent Usability			
	KK	Exp. 1	Exp. 2	Exp. 3
AESTHETICS	.589	.921	.832	.920
DISTANCE	.000	.001	-.042	-.129
KEYPAD TYPE [#]	.730	.671	.751	.760
GROUPING	.075	-.462	-.529	-.667
SEQUENCE 1 [#]	.113	.352	.197	.397
HAND-DOMIN	-.127	-.002	-.125	-.203
SEQUENCE 2	-.306	.233	.137	.153
SAFETY	.137	-.019	-.006	-.061

Table 1. Correlations (bold: $p < .01$) and coefficients of contingency ([#]) of aesthetics and seven inherent usability variables with apparent usability for the experiment in Japan (KK) and for the three experiments in Israel.

Figure 3. Correlations between the design variables and perceived (apparent) usability in the original study and the three replication studies, as reported in Tractinsky (1997).

References

- [1] Kurosu, M., and Kashimura, K. (1995). Apparent Usability vs. Inherent Usability: experimental analysis on the determinants of the apparent usability. Conference Companion on Human factors in Computing Systems, Denver, USA, pp. 292-293.
- [2] Tractinsky, N., (1997). Aesthetics and apparent usability: empirically assessing cultural and methodological issues. ACM CHI Conference Proceedings on Human Factors in Computing Systems (CHI 97), pp. 115-122.