

Understanding, Measuring, and Designing User Experience: The Causal Relationship Between the Aesthetic Quality of Products and User Affect

Haotian Zhou^{1,2} and Xiaolan Fu^{1,*}

¹ State Key Laboratory of Brain and Cognitive Science, Institute of Psychology
Chinese Academy of Science, Beijing 100101, China

{zhouht, fuxl}@psych.ac.cn

² Graduate School, Chinese Academy of Science, Beijing 100101, China

Abstract. This study sought to test the often-taken-granted assumption about the causal relationship between the aesthetic quality of products and user affect by using affective priming paradigm. The results showed that when beautiful web-pages were used as primes, the discrepancy between the response latencies to positive target and negative targets was larger than when the primes were ugly-webpage. A parallel pattern was obtained when pleasant pictures and unpleasant pictures were used as primes. Such findings supported the hypothesis that visual Gestalt of products can lead to affective change independent of reflective beauty judgment. The possibility of employing affective priming procedure to measure product beauty is also discussed in the light of the experiment results.

Keywords: user experience, aesthetics, beauty, affect, affective priming.

1 Introduction

Since the first documented attempt to define user experience (UX) in 1996 [1], a major shift of focus from functionality and usability to non-pragmatic or hedonic aspects of products has been observed in the field of human-computer interaction and interaction design. Available evidence to date all points to the same conclusion that the hedonic aspects of a given interface can significantly influence the user experience of that product [8, 19]. Of all the non-instrumental qualities of products, beauty has been gaining prominence amongst HCI researchers, and striving for beauty has become one of ultimate goals of product design process [4, 16].

Though limited studies have yield important findings attesting to the leverage beauty has on UX [17], the underlying mechanism still remains nebulous [9]. Norman [15] argues that beautiful products induce positive affect in users, which in turn facilitates the user-product interaction process. Though Norman's claim has enormous appeals among researchers [9], it is of both theoretical and practical importance to subject this

* Corresponding author.

causal chain to careful examination. Until recently, much of the emphasis has been put on testing out the link between positive affect and UX [e.g. 13], while the first part of Norman's claim has been largely assumed to be true. For example, Hassenzahl [9] asserts that beauty judgment is driving by the affect evoked by the visual Gestalt of products (Fig. 1a). Yet, given the lack of empirical evidence, one can propose an alternative model in which the direction of causality between beauty judgment and elicited affect reverses (Fig.1b). Note the principal difference between the two models is that whether conscious aesthetic evaluation is the precondition for the visual Gestalt of products to exert impact on users' affective state or not.

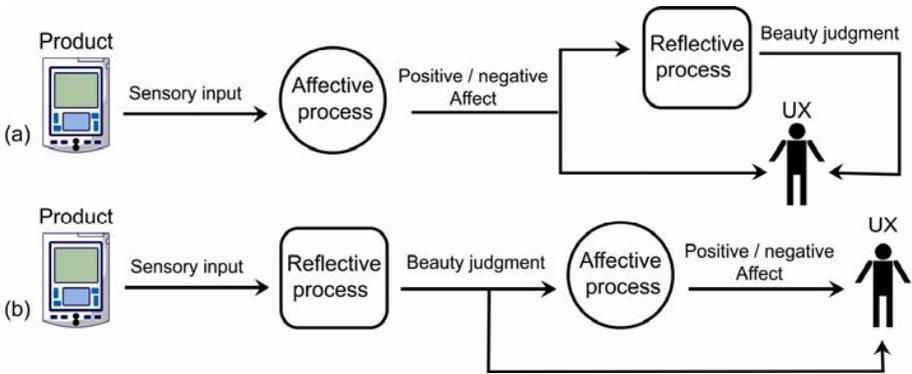


Fig. 1. Outlines of two competing models of how beauty leads to affect: (a) Norman's model; (b) alternative model. Despite their differences, we acknowledge that cognitive appraisal per se is capable of influencing UX (user experience).

Since the rationale behind many UX professionals' advocacy of assigning more weight to aesthetics during product design process rests heavily on Norman's model [15], this growing interest in beauty might be rendered groundless if the alternative model proves to be right (see Discussion section for a detailed explication). In fact, Hassenzahl [9] has indicated the necessity of carefully scrutinizing Norman's claim.

The lack of research to discriminate between the two competing models may be due to the inability of traditional UX methodologies to dissociate affective process and reflective process. In the present research, we endeavored to tackle this issue by adopting an approach often used to investigating the interplay between cognition and emotion. In addition, we also intended to demonstrate the possibility of adapting this paradigm for use as a promising measuring instrument of product beauty.

2 Methods

2.1 Overview

The affective priming paradigm developed by Fazio *et al.* [6] was adopted to assess the automatic affective response evoked by the visual Gestalt of products. The

underpinning of affective priming paradigm is the so-called congruency effect, that is when the affect induced by the prime is of the same valence as target (e.g. both are positive) the evaluation of target valence (i.e. whether it is negative or positive) will be facilitated as compared to the response to a incongruent target (in this case, a negative one). Thus, rather than explicitly inquiring participants about the affect elicited by certain stimuli—primes, it is inferred from participants’ response to another distinct yet affectively related stimuli—targets [10]. Consequently, this paradigm enables us to examine the affective effect of beauty free from possible distortion of conscious reflection (e.g. social desirability).

2.2 Participants

25 undergraduates participated in the experiment (12 men and 13 women), with a mean age of 21.5 years and an age range of 19-23 years. All of them were right-handed and had normal or corrected-to-normal vision ability. All participants were reimbursed upon the completion of the experiment.

2.3 Materials

In this study, we concentrated on the beauty of website specifically. The screen shots of 100 English-language web pages constitute the primary stimulus pool. Among them, 50 are badly designed web pages (Fig. 2a) taken from www.webpagesthatsuck.com and



Fig. 2. Examples of web-pages used in this study: (a) well-designed webpage; (b) bad-looking webpage

other sources; and the remainders are beautifully designed web pages (Fig. 2b) from a few design-award winner lists (e.g. www.worldwidewebawards.net). 30 Chinese undergraduate students (15 males and 15 females), participated in a rating procedure designed by Lindgaard *et al.* [14], in which they were asked to give beauty judgment to all the candidate web-pages presented one by one in random order. Fig. 3 depicts the time course of a single trial in the rating session: each webpage was on screen for 500ms after a 800ms fixation symbol, and then, participants assigned a visual appeal score to that webpage via a sliding bar. Participants were unable to proceed to a second webpage unless they finished rating properly. An average visual attractiveness score

was computed for each webpage following the algorithm suggested by Lindgaard *et al.* [14], and the candidates were ranked accordingly. The 20 most appealing and the 20 least appealing websites were retained for usage in subsequent experiment.

60 affective pictures taken from Native Chinese Affective Picture System [2] served as controls for webpage in subsequent study. The valence score of each picture was assessed using 9-point valence scale with 1 designating extremely unpleasant and 9 extremely pleasant. One third of the pictures are unpleasant ones such as bloody scene (mean rating = 2.18, $SD = 0.20$), 20 are of positive valence such as smiling baby (mean rating = 7.47, $SD = 0.19$), and the remainders are neutral ones such as common tool (mean rating = 4.99, $SD = 0.17$). The targets consist of 20 positive Chinese adjectives (e.g. outstanding) and 20 negative ones (e.g. selfish) selected from the list standardized by Luo and Wang. These words were rated using the same valence scale. The mean valence rating of the 20 negative adjectives is 2.81 ($SD = 0.10$) while that of the 20 positive one is 6.81 ($SD = 0.16$). The average familiarity of positive words is 5.22 ($SD = 0.22$), slightly higher than that of negative ones, 4.73 ($SD = 0.42$).

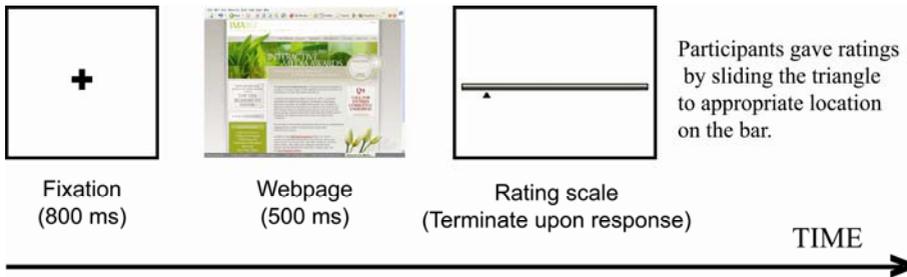


Fig. 3. Time sequence of the webpage rating procedure

2.4 Experiment

The time sequence of the affective priming procedure is shown in Figure 4. Each trial started with a fixation symbol (600 ms) followed by the prime (100 ms). After the prime was a blank screen lasting 50 ms, then the target appeared. Participants were required to judge whether the target was a positive word or a negative one as fast and accurate as possible by pressing specified keys on keyboards. After the judgment, the target disappeared and the program slept for 3 s before next trial resumed. Reaction time and accuracy for each target were recorded.

The 40 web-pages retained plus the 60 affective pictures constituted the prime pool of this experiment. The primes were classified into five categories—beautiful webpage (BW), ugly webpage (UW), pleasant pictures (PP), unpleasant pictures (NP), and neutral/control pictures (CT). The experiment employed a fully crossed 5 (prime categories) by 2 (target valence: negative word vs. positive word) within-subject design. Each of the 40 targets appears five times, once in one of the five prime categories with the stipulation that the same instance from a given prime category can only be paired with one instance from a given target category. Such a pairing scheme

guaranteed that the same set of targets was used as its own control with respect to prime categories. The whole experiment session consists of a training block with 16 trials and 200 experimental trials, which were divided evenly into two blocks with a break in-between.

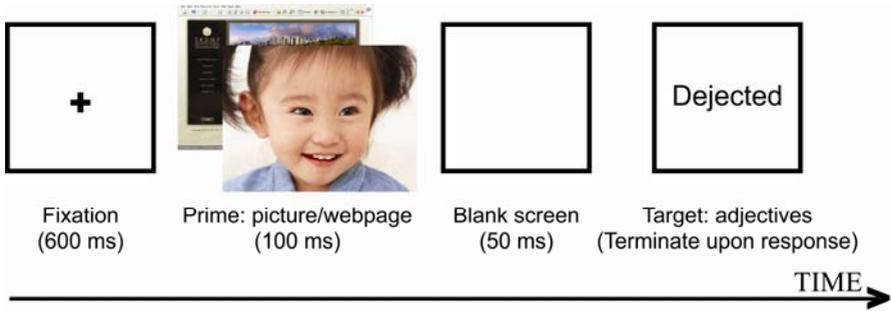


Fig. 4. Time sequence of a single trial in the affective priming procedure

3 Results

3.1 Data Screening and Preliminary Data Analysis

The priming data were screened for outliers by excluding trials with reaction times below 250 ms or above 1000 ms (8.6% of all trials). After correction for outliers, trials with a false response (1.7 % of the remaining trials) were also eliminated from subsequent analysis.

Preliminary analysis shows that response latencies to positive targets ($M=608.9$) were shorter than those to negative words ($M = 632.18$), $F(1, 24) = 17.15$, $p < .001$. Such a positive-target-premium (PTP) is a typical finding in previous affective priming studies [3, 20]. Past affective priming studies [3] have demonstrated that the direction of PTP score variations can be used to infer the valence of the affect elicited by primes. Specifically, PTP score decrease is related with negative affect while its increase with positive affect. Therefore, we investigated priming effect of different types of primes by observing the variations of PTP score as a function of prime categories. PTP score for each prime category was computed subject-wisely by subtracting the mean response latency to positive targets from the mean latency to negative ones, resulting in five PTP scores (associated with each of the five prime categories) per subject.

3.2 Priming Effects of Prime Categories

If beauty judgment is indeed an affect-driven response (i.e. affect change precedes beauty judgment, Fig. 1a), it can be hypothesized that the priming effect (indexed by PTP scores) of visual attractiveness of web-pages (ugly vs. beautiful) should resemble that of pleasantness of pictures (non-pleasant vs. pleasant).

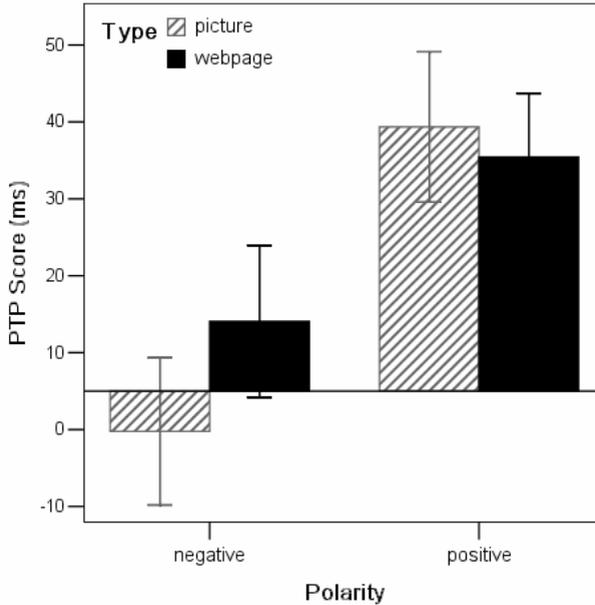


Fig. 5. Mean PTP scores of targets as a function of prime type and prime polarity (Table 1)

To test out this hypothesis, data related with neutral primes were dropped, and the remaining four prime categories used in previous analysis were recoded through two two-level variables—type (webpage and picture) and polarity (negative and positive). Table 1 shows the how different types of primes fall into corresponding cells determined by type and polarity. The PTP score was then analyzed via a 2 by 2 (type by polarity) ANOVA with both variables as repeated measures. The outcome clearly supports the hypothesis (Fig. 5). Of the two main effects and one two way interaction, only the main effect of polarity reaches significance, $F(1, 24) = 32.37, p < .001$, with PTP score related with positive primes ($M = 37.39, SD = 8.00$) much higher than that related with negative primes ($M = 6.92, SD = 8.71$).

Table 1. Correspondence between prime categories and their values on polarity and type

		Polarity	
		negative	positive
Type	picture	NP	PP
	webpage	UW	BW

Post hoc comparisons shows that priming with both ugly web-pages and non-pleasant picture led to significantly decreased PTP score compared with neutral primes: both $t(24) > 2.30; ps < .03$. On the other hand, when the priming effects of

beautiful web-pages and pleasant pictures were compared with neutral primes respectively, PTP scores increased as predicted. However, neither increase reached significance, $ps > .45$ (a detailed account for this unexpected finding as well as its implication is provided in the discussion section). The means and standard deviations for PTP scores associated with all five prime categories are displayed in Table 2.

Table 2. Means and standard deviations for PTP score as a function of prime categories

Prime type	Mean (ms)	Std Deviation (ms)
BW	35.43	41.50
UW	14.06	49.37
PP	39.35	48.90
NP	-.23	47.82
CT	32.79	47.24

4 Discussion

4.1 Beauty and Affect

In the present study, we found that ugly web-pages' influence on subsequent adjective-evaluation task is similar to that of non-pleasant pictures, while beautiful web-pages similar to pleasant pictures. It can be inferred by analogy that ugly web-pages induced negative affect in participants; whereas beautiful web-pages pushed the affective state toward the positive direction. In the present study, participants were required to concentrate on target evaluation tasks rather than the primes. With little attention being directed to prime, any affective reaction evoked by the prime was unlikely to be contaminated by conscious reflective processing, and thus can be seen as the direct outcome of visual Gestalt's impact on affective system (Fig. 1a). Such findings bear out the often-taken-for-granted assumption about the linkage between beauty and affect—visual Gestalt of products is capable of changing affective state of users independent of explicit judgment about product aesthetics.

What would be the consequences if the alternative model outlined in Fig. 1b were true? According to this competing model, visual Gestalt of a product cannot impact users' affective state until users make a conscious aesthetic judgment about its appearance. Yet, in most goal-mode (i.e. driven by predetermined goal) [7] user-product interaction context, it seldom happens that user will explicitly evaluate the visual appeal of the product before proceeding to interact with it. Accordingly, at least in the case of goal-mode products; much of the discussion about making beauty a design goal would become irrelevant, because in this model, the involvement of conscious reflection is required for beauty to influence UX (Fig. 1b).

4.2 Direct Measure vs. Indirect Measure

Affective priming paradigm has recently been developed as an indirect measure of stimuli (e.g. food or friends) valence to replace the more conventional direct measures (e.g. questionnaire) in the field of social psychology [5]. The edge of indirect measure over direct measure is that the former is less susceptible to distortions and bias by respondents (e.g. impression management). However, UX field has been slow in taking up this recent advancement in measuring methodology and traditional measuring instruments (e.g. Likert scales) still predominate [12].

In fact, the unexpected result of our study testifies to the possible caveats associated with the direct measurement. Recall that, compared to neutral pictures, neither beautiful web-pages nor the pleasant pictures were capable of increasing PTP scores significantly as predicated. There are two possible explanations for this finding. The first one is that both pleasant pictures and beautiful web-pages failed to alter affective states of participants, assuming that neutral pictures did not induce affective changes. However, this account can hardly reconcile with two facts (1) past affective priming studies have consistently demonstrate that positive pictures can induce affective reaction [e.g.11, 19] and (2) in this study, negative primes resulted in affect change as indexed by significant PTP score decrease. Hence, we believe the alternative account makes more sense—neutral pictures led to positive affect in participants. After reexamining the neutral pictures, we noticed that most of the pictures are beautifully designed artifacts (e.g. antique) or pleasing geometric patterns (Fig. 6). Their neutrality on rating scale is likely due to the distortion of raters' cognitive process. Despite the initial pleasant affect brought about by the visual Gestalt of a streamlined hairdryer, the raters might have engaged in a reflective process (e.g. how can such mundane object make me feel good?) which dismissed this affective change; and therefore rated it as emotionally bland. Recently, Zhou *et al.* [21] discovered that Chinese participant took the neutral schematic face (explicitly rated as such) as positive one in an implicit task (i.e. categorization). Admittedly, including an extra condition, where no primes precede targets, to our experiment would help clarify this ambiguity. Nonetheless, the present study points to the precaution one should take when interpreting the results of beauty studies employing direct measuring instruments.

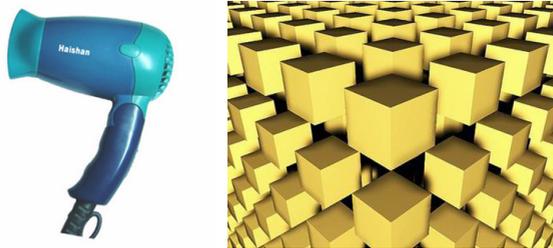


Fig. 6. Examples of emotionally-neutral pictures used in the experiment

Notwithstanding the uncertainty concerning neutral pictures, the present research unequivocally demonstrated that PTP score was capable of distinguishing good designs from bad ones. Therefore, we provided some empirical evidence suggesting the possibility of using affective priming paradigm can be successfully as a promising alternative measuring instrument in future UX research.

5 Conclusion

By differentiating between two opposing accounts of how beauty creates affect (Fig. 1), the present study demonstrated that product beauty is one of the ‘many circumstances in which affective reaction precedes the very cognitive appraisal on the which the affective reaction is presumed to be based [18]’. Note that in our experiment, the affective changes of participants were inferred from PTP scores; therefore, such evidence is indecisive at best. Clearly, more research employing alternative procedures such as physiological instrument are needed to validate our conclusion.

Showing that beautiful and ugly products can be set apart on their effect on PTP scores, this study has interesting implication for the important question of how to measure beauty [9]. Though we speculated about the possibility of adopting affective priming paradigm to measure visual attractiveness of products, the adjustment necessary for achieving this end has yet to be specified.

Acknowledgments. This research was supported by grants from 973 Program of Chinese Ministry of Science and Technology (#2006CB303101), and the National Natural Science Foundation of China (#60433030).

References

1. Alben, L.: Quality Of Experience: Defining the Criteria for Effective Interaction Design. *Interactions* 3, 11–15 (1996)
2. Bai, L., Ma, H., Huang, Y.X., Luo, Y.J.: The Development of Native Chinese Affective Picture System-A Pretest in 46 College Students. *Chinese Mental Health Journal* 19, 719–722 (2005)
3. Banse, R.: Affective Priming with Liked and Disliked Persons: Prime Visibility Determines Congruency and Incongruency Effects. *Cognition & Emotion* 15, 501–520 (2001)
4. Desmet, P.M.A., Hekkert, P.: The Basis of Product Emotions. In: Green, W., Jordan, P. (eds.) *Pleasure with Products, beyond Usability*, pp. 60–68. Taylor & Francis, London (2002)
5. Fazio, R.H., Olson, M.A.: Implicit Measures in Social Cognition Research: Their Meaning and Use. *Annual Review of Psychology*, pp. 297–328 (2003)
6. Fazio, R.H., Sanbonmatsu, D.M., Powell, M.C., Kardes, F.R.: On the Automatic Activation of attitudes. *Journal of Personality and Social Psychology* 50, 229–238 (1986)
7. Hassenzahl, M.: The Thing and I: Understanding the Relationship between User and Product. In: Blythe, M., Overbeeke, K., Monk, A., Wright, P. (eds.) *Funology: From Usability to Enjoyment*, pp. 31–42. Kluwer Academic Publishers, Dordrecht Boston London (2003)

8. Hassenzahl, M.: The Interplay of Beauty, Goodness, and Usability in Interactive Products. *Human-Computer Interaction* 19, 319–349 (2004)
9. Hassenzahl, M.: Aesthetics in Interactive Products: Correlates and Consequences of Beauty. In: Schifferstein, H.N.J., Hekkert, P. (eds.) *Product Experience*, Elsevier, Amsterdam (2006)
10. Hermans, D., Baeyens, F., Lamote, S., Spruyt, A., Eelen, P.: Affective Priming as an Indirect Measure of Food Preferences Acquired through Odor Conditioning. *Exp. Psychol* 52, 180–186 (2005)
11. Hermans, D., Spruyt, A., De Houwer, J., Eelen, P.: Affective Priming with Subliminally Presented Pictures. *Can J Exp Psychol*. 57, 97–114 (2003)
12. Kuniavsky, M.: *Observing the User Experience: A Practitioner's Guide to User Research*. Morgan Kaufmann, San Francisco (2003)
13. Lyubomirsky, S., King, L., Diener, E.: The Benefits of Frequent Positive Affect: Does Happiness Lead to Success. *Psychological Bulletin* 131, 803–851 (2005)
14. Lindgaard, G., Fernandes, G., Dudek, C., Brown, J.: Attention Web Designers: You Have 50 Milliseconds to Make a Good First Impression! *Behaviour & Information Technology* 25, 115–126 (2006)
15. Norman, D.A.: *Emotional Design: Why We Love (Or Hate) Everyday Things*. Basic Books, New York (2004)
16. Norman, D.A.: Introduction to This Special Section on Beauty, Goodness, and Usability. *Human-Computer Interaction* 19, 311–318 (2004)
17. Tractinsky, N., Katz, A.S., Ikar, D.: What is Beautiful is Usable. *Interacting with Computers* 13, 127–145 (2000)
18. Zajonc, R.B., Markus, H.: Affective and Cognitive Factors in Preferences. *The Journal of Consumer Research* 9, 123–131 (1982)
19. Zhang, P., Li, N.: The Importance of Affective Quality. *Communications of the ACM* 48, 105–108 (2005)
20. Zhang, Q., Li, X.: Affective Priming Effects under Two SOA Conditions. *Chinese Journal of Applied Psychology* 11, 154–159 (2005)
21. Zhou, G., Fu, X., Hayward, W.G., Locke, V., Pellicano, E.: Cultural Difference in the Application of the Diagnosticity Principle to Schematic Faces. *Journal of Cognition and Culture* 5(1), 240–247 (2005)