

# **User Satisfaction, Aesthetics and Usability**

## *Beyond Reductionism*

Gitte Lindgaard & Cathy Dudek

*Carleton Human Computer Interaction Institute*  
*Carleton University, Ottawa, Ontario, Canada*

[gitte\\_lindgaard@carleton.ca](mailto:gitte_lindgaard@carleton.ca) , [cdudek@chat.carleton.ca](mailto:cdudek@chat.carleton.ca)

**Abstract:** Results from a series of web site studies suggest that the concept of user satisfaction comprises more than perceived aesthetics and usability. Satisfaction was repeatedly found to be a complex construct comprising 'emotion', 'likeability', and 'expectation' as well. A web site very high in appeal but low in usability scored highly on user satisfaction when first encountered. However, when faced with serious problems in a usability test, users' overall level of satisfaction dropped considerably, but perceived aesthetics remained unchanged. Given the known importance of the first impression for subsequent judgments, our results suggest that user interface designers of e-commerce sites would be well advised to design pretty and usable sites. Designing for user efficiency and effectiveness alone is not enough unless the products and services offered on a web site are unique in the world.

**Key words:** satisfaction, aesthetics, appeal, usability, emotion

## **1. INTRODUCTION**

According to the ISO 9241-11 standard, user satisfaction is supposed to contribute to usability along with effectiveness and efficiency (ISO, 1997). Among the plethora of usability assessment techniques, with few exceptions (Kirakowski, 1996), hardly any concern, or include, measures of user satisfaction. This is understandable when the goal is to make users more efficient and effective. Indeed, many measurements of user satisfaction tend to be limited to assessing "what users think of [a given application]" (Macleod, Bowden, Bevan & Curson, 1997). Fewer still are concerned with the emotional impact of an interface (Kim & Moon, 1998). Apparently, it is

assumed that users will like and accept a highly usable application that enables them to do their job quickly and efficiently. This is a form of usability reductionism, where joy (or even satisfaction) is merely a by-product of great usability (Hassenzahl, Beau & Burmester, 2001). The assumption that productivity enhancement automatically fosters satisfaction may be justified in traditional office applications where a person's livelihood, a company's profit, public well-being or safety may depend on just that. However, even if we accept that assertion, it still does not follow that satisfaction is a component of usability – satisfaction may result primarily from usability issues.

In this paper we argue that satisfaction may be a by-product of great usability in traditional office environments, and that satisfaction can be defined in terms of efficiency and effectiveness. However, on the World Wide Web where users choose to spend their leisure time finding information, seeking entertainment, or shopping, and where the next competing site is but a click away, we suspect that users employ quite different criteria in evaluating their experience. We also believe that this evaluation depends upon users' needs and goals. In order to widen the notion of user satisfaction beyond efficiency and effectiveness of the user experience, researchers must start to think of usability as part of a satisfying user experience. In exploring the satisfaction construct, we consulted research in Human Computer Interaction (HCI), neurophysiology, and marketing.

Research in the consumer and marketing literature has shown that consumers readily recall the emotional content of customer service encounters and that they use semantically different words to describe their experiences with different industry sectors (Edwardson, 1998). The literature has also found that 'satisfied' customers are just as likely to defect as those who are neutral or mildly dissatisfied (Jones & Sasser, 1995). That is, unless customers are 'highly satisfied' with a company's goods and services, the company cannot take customer loyalty for granted. If user satisfaction is motivated by different criteria and if the questions phrased in 5-point or 7-point scales asking them to judge 'appeal'/'attraction' or 'pleasantness' of the interactive experience fail to capture the essence of user satisfaction in a given context, then we may be misled in our interpretation of satisfaction scores. The research we report here is motivated by a need to 'unpack' the notion of user satisfaction in the context of e-commerce web sites by listening to what users tell us about their interactive experience. Increasing attention to user satisfaction, however, does not mean that we can afford to neglect the performance-related aspects of usability that we have traditionally measured. We also need to learn how satisfaction relates to user effectiveness and efficiency. User Interface (UI) designers working within

strict budgetary and time constraints need guidance on how best to divide resources to satisfy both and balance aesthetics with efficiency factors in their designs. As well as understanding how to measure user satisfaction, our aim is therefore ultimately to derive valid design guidelines for a wide range of interactive technologies.

Our concern with satisfaction arises from very robust findings in the neurophysiological literature where researchers constantly find that emotional responses are strikingly immediate, occurring within 3-4 milliseconds of a stimulus being shown (Bornstein, 1992; Zajonc, 1980). Thus, according to this research, emotional responses are pre-attentive and precede cognitive ones. The implications for web design are obvious and pervasive: if users decide that they dislike what they see in less than five milliseconds, then they may click onto the next site even before they have taken in any information it offers. The strength of the 'first impression', characterized by what psychologists call a 'primacy effect', has long featured very prominently in the psychological literature, (Anderson, 1981; Anderson, 1982) even in areas involving expert judgement such as diagnostic medicine (Lindgaard, 1985). Basically, judgments are overwhelmingly based on the first impression. Where a primacy effect occurs, the stimulus presented or detected first receives a disproportionate amount of attention. The subsequent search for evidence to substantiate the judgment already made is biased in favour of searching exclusively for confirmatory evidence while ignoring contradictory evidence, giving rise to the so-called 'confirmation bias' (Mynatt, Doherty & Tweeney, 1977). So, if users have already decided they dislike a site, they will interpret virtually all information as being more negative than if their first impression were positive.

A recent study by Tractinsky, Katz & Ikar (2000) investigated the extent to which "the initial perceptions of aesthetics-usability relationships hold after a period of system use, and whether these perceptions are affected by the interface's perceived aesthetics and/or by the actual usability of the system" (p.131). They found that judgments of interface aesthetics were not affected by traditional usability factors. Using an array of ATM interfaces with identical content but varying in layout and in appeal, as determined in the pre-experimental phase, they introduced several usability problems to which the 'low usability' but not the 'high-usability' group was exposed. While this manipulation affected the task-completion times, judgments of interface aesthetics did not change and neither did judgements of perceived usability. The researchers concluded that the relationship between usability and satisfaction is not orthogonal, which lead to their provocative claim that "what is beautiful is usable". Their findings suggest that the two concepts are correlated. One may regard this as a form of design reductionism, where

joy of use can be brought about by aesthetics alone, even in the face of usability problems (Hassenzahl et al, 2001).

By contrast, studies performed in our laboratory suggest that usability may be judged independently of interface aesthetics when users are confronted with severe usability problems (Lindgaard & Dudek, 2002). For example, one site perceived to be extremely high in aesthetics but very low in usability scored substantially lower on usability, but as highly on satisfaction as another site perceived to be high in both usability and aesthetics. However, Tractinsky et al.'s (2000) subjects undertook tasks as in a traditional usability test, our results were based on retrospective self reports obtained in interviews immediately after subjects had inspected the site for 10 minutes. Subjects were instructed to verbalize their experience of the interaction. In the absence of a requirement to complete a set of usability tasks designed to test the extent to which usability flaws get in the user's way, it is quite possible that our subjects' experiential descriptions were based entirely on the first impression of the site. This begs the question of the strength and duration of the first impression in a web environment.

Much of the psychological literature on the primacy effect suggests that it is paramount and may determine any judgmental outcome (Anderson, 1982; Slovic & Lichtenstein, 1971), presumably including a judgment of 'satisfaction'. By the same token, the so-called 'mere exposure effect' suggests that the strength of the initial (emotional) impression starts to wane once the exposure time exceeds 50 milliseconds (Bornstein, 1992). A strongly negative first impression could be commercially damaging to a company aiming to increase its online sales, particularly if it competes with many others offering the same goods and services. To ensure a positive first impression, a greater proportion of the usually limited UI design resources may thus have to be devoted to interface aesthetics, perhaps even at the expense of some usability factors. If, however, the emotional first impression does fade as quickly as research into the mere exposure effect suggests, site visitors may well hang around long enough to consider the merits of the information/services/products offered on the site even if the first impression does not evoke a 'wow' effect. In that case, UI designers would be quite justified in their continued quest to create usable sites at the expense of making them strikingly 'beautiful'. In Tractinsky et al.'s (2000) study, users did not change their mind after completing tasks in which they encountered usability problems. The experiment reported here was designed to test the robustness of these authors' finding when users are exposed to more serious usability problems.

## **2. FRAMEWORK OF THE STUDY**

A web site offering exclusive writing utensils was used. In an earlier experiment, this site was found to be significantly higher in appeal and lower in perceived usability than several other sites using the same method of investigation and different groups of subjects (Lindgaard & Dudek, 2002). In contrast to several of the other sites employed in the same series of experiments, all of which were typical shopping sites, the pen site contained no prices on goods or monetary transaction modules. Its purpose was thus apparently to market rather than sell goods. The present study proceeded in two phases. In Phase 1, an heuristic evaluation was conducted to identify the nature, location, and severity of usability problems in the web site. The outcome of this evaluation served as a basis for selecting and designing user tasks to be performed in the subsequent usability test. Phase 2 comprised the usability test enabling a comparison with an earlier study of the same web site, as follows:

Present study: Browse site (10 min) → unstructured interview → usability test → unstructured interview

Previous study: Browse site (10 min) → unstructured interview

### **2.1 Heuristic Evaluation**

The heuristic evaluation revealed some 157 instances of moderate ( $n = 45$ ) to severe ( $n = 112$ ) usability problems. By our definition, a moderate problem gets in the user's way but does not prevent progress towards accomplishing a goal, whereas a severe usability problem does. As can be seen in Table 1 below, the majority of the problems involved navigation ( $n = 77$ ). Some 29 of these concerned hyperlinks (e.g. looks like a hyperlink but is not; does not look like a hyperlink but it is; does not behave like a hyperlink). The remaining 48 were due to unusual, unpredictable, inconsistent navigation rules or awkward navigation operations. For example, menus were 'floating' in and out of view, forcing the user to select an option very quickly, and the active area surrounding a point on a map was tiny, requiring such fine motor movements that it was almost impossible to point precisely. Another group of problems ( $n = 33$ ) concerned visibility and comprehension of text/objects, being marred by confusing or misleading vocabulary, contrast problems, or text displays that were partly obscured by overlapping graphics. Some ( $n = 24$ ) were due to inconsistent system behaviour or display rules, e.g. information displayed in the topmost screen position in one screen was at the bottom of the next, or text would move while the user was reading it. The remainder ( $n = 23$ ) were unique problems not falling into any of the above categories.

Nature of problem	N
Navigation, not including hyperlinks	48
Hyperlinks	29
Visibility & comprehension	33
System behaviour/display rules	24
Other	23
<b>Total</b>	<b>157</b>

*Table 1.* Number of usability problems by category

## 2.2 Selection of User Tasks

Upon loading, the site played an animated introduction taking several minutes with accompanying soft music matching the rhythm of the colourful, entertaining animation. This introduction looked more like a TV commercial than an e-commerce site. The site comprised three main sections, each of which was sub-divided into several sub-sections with these sub-dividing further to a maximum of four levels. The eight user tasks were selected to satisfy two objectives: (1) all three main sections were represented and, (2) they represented different levels of difficulty and number of clicks to the target information. Tasks appeared straight-forward, for example, asking subjects to find out when the company was established, whether the company had any job openings, and to find the nearest retail outlet.

## 2.3 Procedure

Twenty subjects, 10 males and 10 females, were recruited from around the University Campus in a semi-random fashion, ensuring that English was their first language, that they were regular Internet users (2-10hrs/week), and that they had no UI design or evaluation experience. Subjects first inspected the site for 10 min, having been told to concentrate on their interactive experience and pretending they were looking for a gift for a special person to send as an apology. At the end of the 10 minutes, an unstructured interview was conducted to elicit as many experience-related statements as possible. Next, they completed the eight usability tasks, given to all subjects in the same sequence, beginning with the easy tasks, and ending with more difficult tasks. This was done to give subjects a sense of success and motivate them to work through all the tasks, as we expected some of the searches to be unsuccessful. All subjects attempted all tasks, and there were no time- or accuracy constraints. Subjects were allowed to give up if they

were unable to retrieve the information needed to complete the task. At the end of the usability test another interview was conducted in the same manner as before to learn whether the initial impression remained constant or whether appeal was attenuated compared with the first interview. Subjects were tested individually in sessions lasting up to 1.5 hours. They were paid \$15 Canadian for their time.

## **2.4 Data Analysis**

Interviews were audio taped. Data were transcribed ad verbatim and submitted to a content analysis. Statements were divided into 5 categories: aesthetics, emotion, expectation, likeability and usability. Aesthetics-type statements all referred to visual qualities of the interface (too much blue, too much white space, bright, pretty, pleasing to the eye). Emotion statements, defined in terms of Russell's Circumplex Model of Affect (Russell, 1980), were those that could finish the sentence "It was [made me feel]...." (uplifting, relaxing, calming, frustrating) and represented a concept that could reasonably have been contained in Russell's (1980) Model. Expectation statements expressed thoughts about components that the subject was surprised to find in the interface, thought would be there, or should have been there. In some cases, they used the word 'expected' in the statement; in others their expectation could be derived from what was said. For example, "I would have thought they would put all the pens together" reflected an unmet expectation. Likeability statements were overall judgements about the site or comparisons with other sites (better than, I like it, it's okay, fine, not as good as). Finally, usability comments were those that referred directly to efficiency or effectiveness, for example, 'there is no back button'; 'I could not click on that', and 'the choices are not logically displayed'. Statements in which subjects merely read aloud screen content were eliminated. Statements were counted once only regardless of the number of times a given word was repeated in an interview.

## **3. RESULTS**

In this section, the data from the first interview (before usability test - Experiment 2) are compared with those obtained in the earlier experiment (browsing only - Experiment 1), details of which are reported elsewhere (Lindgaard & Dudek, 2002). The data comparing the 'before' and 'after usability test' conditions are reported in section 3.2. First, we wanted to explore whether people would have more or less to say about their experience when they expected to complete a usability test, so a two-tailed t-

test was performed for the total number of statements. Participants had more to say when they were expecting to complete a usability test (Experiment 2) than when they were just browsing the site (Experiment 1) with no particular purpose in mind. ( $t(38)=-2.24$ ,  $p<.05$ ), as shown in Figure 1. Looking next to see if the strength or quality of the experience differed, the proportion of positive statements each group made are displayed in Figure 2.

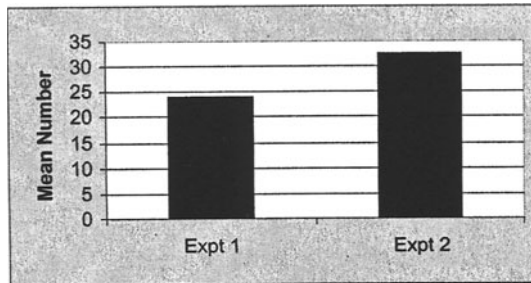


Figure 1. Mean number of all statements.

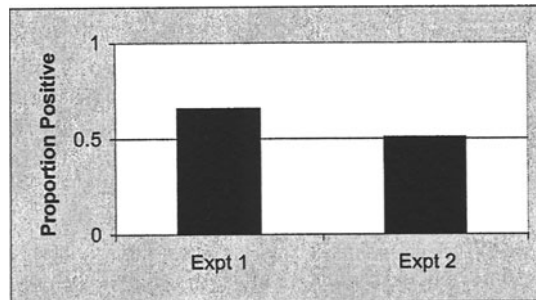


Figure 2. Proportions of all types of statements.

A two-tailed t-test comparing the proportion of positive statements in both experiments showed that the experience of 'just browsing' subjects was more positive than that for subjects who were preparing to complete a usability test ( $t(38)= 1.99$ ,  $p<.05$ ). Observation of subjects during the browsing session in both experiments showed that 'browsing only' subjects tended to move around the site in an ad hoc fashion. By contrast, those who knew the usability tasks would follow set out systematically to look through as much of the site as they could in the 10min browsing time. These latter subjects saw more of the usability problems owing to the sheer difficulty of navigating the site. This difference in browsing behaviour probably accounts for the difference in positive statements.

Consistent with earlier findings, statements fell into the categories mentioned earlier (likeability, emotion, aesthetics expectation and usability).



The mean number of statements in each category differed between experiments as shown in Figure 3. These results show that people had more to say in the second experiment than the first, particularly about likeability ( $t(32)=-2.50$ ,  $p<.05$ ), and aesthetics ( $t(26)=-2.75$ ,  $p<.01$ ). The mean number of emotion, expectation and usability statements did not differ ( $p>.05$ ) from the first to the second experiment.

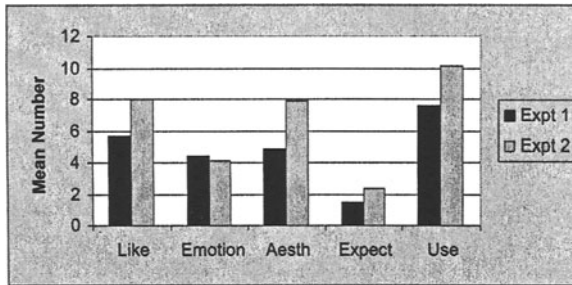


Figure 3. Mean number of all types of statements.

### 3.1 Usability Test

Of the 160 tasks attempted across all subjects, only 79 (49.38%) were completed successfully. The average success rate was 3.95 of the eight tasks. No one completed all eight tasks successfully, and none of the tasks were completed successfully by all subjects. Only one was completed successfully by more than 15 subjects. Three tasks were completed successfully by 10-15 subjects, and another three by 5-10 subjects. The remaining task was not completed successfully by a single subject. The number of clicks to success exceeded the optimal number in all tasks, ranging from 100% to 2800% above the optimum for successfully, and from 100% to 8650% for unsuccessfully completed tasks. Thus, when subjects did give up, it was only after trying hard. This allowed us to conclude that they took the tasks seriously. These data also confirm that the level of usability was very low.

### 3.2 Comparison of First and Second Interview

Now comparing the 'before' and 'after usability test' interviews, the data suggest that the user experience changed from the first to the second interview. Although we expected the number of statements to be lower in the second interview because the subjects had already said all that they wanted in the first interview, a one-tailed t-test for paired samples showed

that the total number of statements did not differ ( $p > .05$ ). This is shown in Figure 4.

As before, the proportion of positive statements was calculated. These are shown in Figure 5. A two-tailed t-test for paired samples showed that participants made significantly fewer positive statements after the usability test ( $t(19) = 4.60, p < .001$ ). Thus, although the total number of statements did not change from the first interview to the next, the user experience was apparently more negative after than before the usability test.

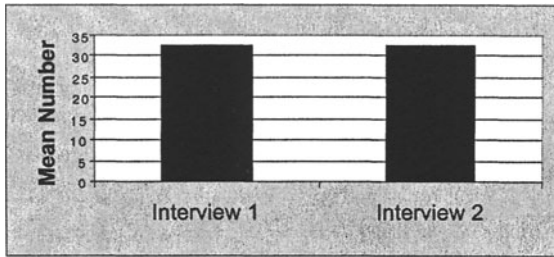


Figure 4. Mean number of all statements.

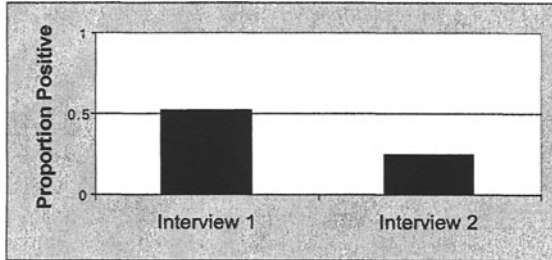


Figure 5. Proportion of positive statements.

Breaking down the statements by category as before showed that there were dramatic changes in some of these. This is shown in Figure 6.

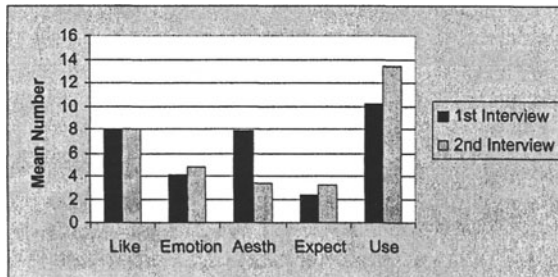


Figure 6. Mean number of all types of statements.

In the second interview subjects tended to say less about aesthetics ( $t(19)=4.57$ ,  $p<.001$ ) and more about usability ( $t(19)=-2.77$ ,  $p<.01$ ). There were no difference the numbers of likeability, emotion and expectation statements ( $p>.05$ ). This is not surprising. One would expect that participants already said what they wanted in the first interview, and there is no reason to think that they would repeat themselves a second time. The larger number of usability statements can be attributed to a greater awareness of usability issues resulting from the task requirements.

A comparison of the proportion of positive statements in the two interviews shows that there was no overall decline in the user experience from one interview to the next in terms of aesthetics ( $p>.05$ ), as shown in Figure 7. Thus, although subjects had less to say about aesthetics in the second interview than in the first, they did not find the site uglier after the usability test.

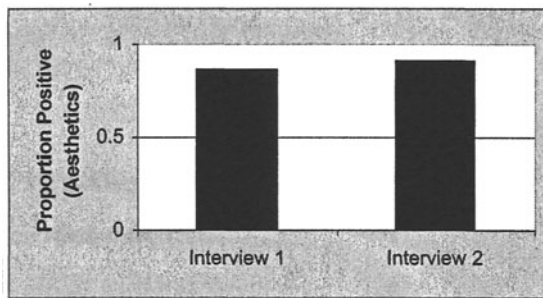


Figure 7. Proportion of positive aesthetics statements.

The pattern of results for the other categories are shown in Figures 8-11. For likeability (Figure 8) and expectation (Figure 9) the overall number of positive statements remained similar ( $p>.05$ ).

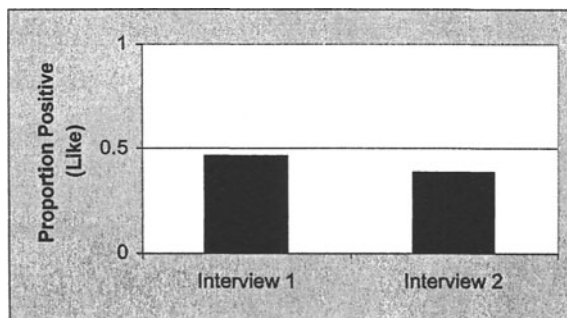


Figure 8. Proportion of positive likeability statements.

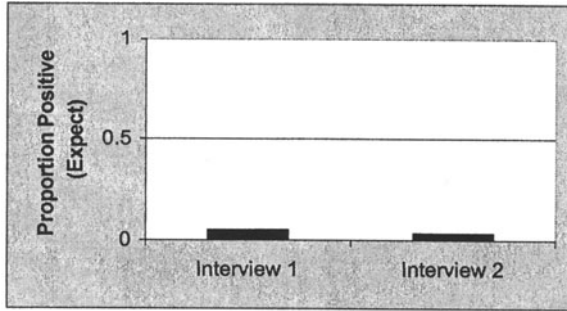


Figure 9. Proportion of positive expectation statements.

For emotion (Figure 10), a two tailed t-test ( $t(18)=4.17$ ,  $p<.001$ ) showed that significantly fewer positive emotions were expressed in the second interview than the first. It appears that participants changed their minds about their experience after encountering usability problems.

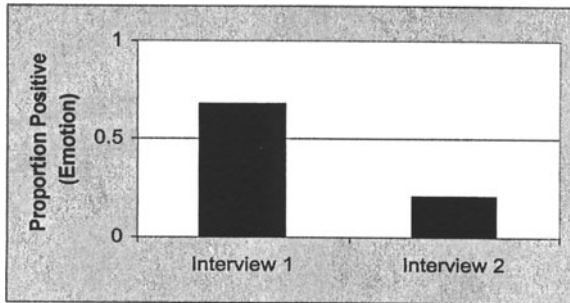


Figure 10. Proportion of positive emotion statements.

For usability statements, there was also a difference. As shown previously in Figure 5, there were more usability statements made in the second interview than there were in the first. However, looking at the proportions of positive and negative statements shows that they changed their minds about their experience regarding usability as well, as shown in Figure 11. A two-tailed t-test for paired samples showed that there were significantly fewer positive usability statements made during the second interview than the first ( $t(19)=2.98$ ,  $p<.001$ ). Thus, although subjects did identify numerous negative usability issues before they completed the usability tasks, these assumed more prominence during the usability test and in the subsequent interview.

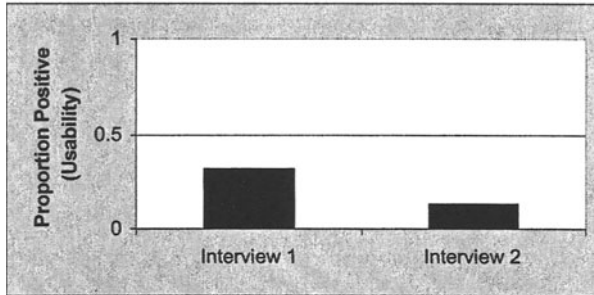


Figure 11. Proportion of positive usability statements.

#### 4. DISCUSSION

Considering first the notion of aesthetics, our results appear, on the surface, to agree with Tractinsky et al.'s (2000) findings in the sense that subjects did not consider the user interface uglier after completing usability tasks than before. However, Tractinsky and his colleagues use the term 'aesthetics' interchangeably with 'affect'. They appear to believe that aesthetics ratings are indicative of user satisfaction. Their subjects rated three aspects of the user interfaces, namely aesthetics, ease of use, and amount of information on the screen. Yet, the authors argue that "there are strong correlations between users' satisfaction from using the system and their perception of its aesthetics and usability" (p. 141). By contrast, our results suggest that the interactive experience comprises at least the five dimensions discussed here, including perceived aesthetics and usability. This repeated finding leads us to argue that usability rightfully belongs under the umbrella of satisfaction and that the notion of user satisfaction is more complex than a correlation between aesthetics and usability. Indeed, our results suggest that the two are not correlated at all: perceived usability is likely to change after encountering usability problems whereas perceived aesthetics is not.

A closer look at Tractinsky et al.'s (2000) findings suggests that the usability problems they introduced did not seriously hamper subjects' performance. The authors describe three usability problems, all of which delayed performance, for example, introducing a delay of nine seconds on average per task. However, all subjects completed all the 11 tasks successfully. By contrast, our subjects completed roughly one half of the eight tasks successfully, as discussed earlier. Our subjects liked the site less overall after experiencing serious usability problems, and the proportion of negative usability comments increased in the before-after comparison of

usability statements. The problem seems to be in the definition of 'aesthetics'. Tractinsky et al. (2000) uses it interchangeably with 'appeal', and 'beauty'. Our results have consistently shown that 'appeal' or what we call 'user satisfaction' comprises more than 'beauty', which by our definition is taken to equate 'aesthetics'. More research is needed to clarify these issues and sharpen the terminology we use to capture and describe the user experience.

With respect to the strength of the first impression, our results suggest that subjects who knew they would be performing usability tasks liked the site less on first encounter than subjects who were 'just browsing'. Thus, the different task demands resulted in different browsing patterns that called more attention to usability. Subjects who knew that they would be asked to complete usability tasks browsed the site in a systematic, goal-oriented fashion, whereas those who were 'just browsing' let themselves be carried away by the show unfolding before their eyes. This raises the issue of site design vis à vis the purpose a given site is intended to fulfil. Because the pen site did not allow users to select and buy items, we believe it was designed to yield a pleasant, but passive experience. In contrast, shopping sites aim to engage users actively and guide them effortlessly through a purchasing transaction. Usability was clearly not a major design objective. However, our usability task demands led subjects to focus on usability. Consequently, the first impression suffered. The first impression would thus appear to depend upon the user's goal: if seeking an obligation-free entertaining experience, subjects pay more attention to the experiential aspects than to usability factors, but when visiting a site to buy goods, the reverse seems to be the case. Thus, user satisfaction seems to be driven by the users' motivation for visiting the site and cannot be reduced to a by-product of aesthetics, usability or even a combination of both.

## **5. CONCLUSION**

The present study suggests that goals determine the users' frame of mind, within which the site is perceived and interpreted, and that this first impression may change as a consequence of facing serious usability problems. While subjects are aware of usability problems even when they are 'just browsing' these affect their opinion of the site less. Finally, the results suggest that concern for traditional usability issues is an integral part of the interactive user experience, however, user satisfaction is a complex construct involving more than an impression of 'aesthetics' or 'usability' alone. The relationship between appeal and user satisfaction, and between perceived/actual usability remain evasive and need much more research to

be clearly understood. While the first impression may be strong and may relate to the immediate appeal of the web site, satisfaction may change as a function of encountering serious usability problems in the context of accomplishing a specific goal. Thus, if it is true, that the first impression is based on immediate appeal, UI designers would be well advised to create aesthetically appealing sites that clearly and immediately reflect its purpose.

## 6. REFERENCES

- Anderson, N.H. (1981), *Foundations of information integration theory*, Academic Press, London.
- Anderson, N.H. (1982), *Methods of information integration theory*, Academic Press, London.
- Bornstein, R. (1992), *Subliminal Mere Exposure Effect, Perception Without Awareness: Cognitive, Clinical, and Social Perspectives*, Bornstein, F. and Pittman, T. (eds) The Guilford Press: New York.
- Edwardson, M. (1998), Measuring Consumer Emotions in Service Encounters: An Exploratory Analysis, *Australian Journal of Market Research*, 6, 2, July, p. 34-48.
- Hassenzahl, M., Beau, A., & Burmester, M., Engineering Joy, (2001), *IEEE Software*, January-February.
- ISO (1997), ISO/DIS 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs): Guidance on usability.
- Jones, T. & Sasser, E. Jr. (1995), Why Satisfied Customers Defect, *Harvard Business Review*, November - December.
- Kim, J. & Moon, J. (1998), Designing towards emotional usability in customer interfaces- trustworthiness of cyber-banking system interfaces, *Interacting with Computers*, 10, 1-29.
- Kirakowski, J. (1996), The software usability measurement inventory: Background and usage, in P. Jordan, B. Thomas & B. Weerdmeester (Eds), *Usability evaluation in industry*, Taylor & Francis, London.
- Lindgaard, G. (1985), Weighting of individuating information elements and base rate in a nursing decision making task involving non-diagnostic case information, Unpublished Masters Thesis, Department of Psychology Monash University, Clayton, Australia.
- Lindgaard, G. & Dudek, C. (2002), What is this beast we call user satisfaction?, submitted to *Interacting with Computers*.
- Macleod, M., Bowden, R., Bevan, N., & Curson, I. (1997), The MUSiC performance measurement method, *Behaviour and Information Technology*, 16, 1-27.
- Mynatt, C.R., Doherty, M.E. & Tweney, R.D. (1977), Confirmation bias in a simulated research environment: An experimental study of scientific inference, *Quarterly Journal of Experimental Psychology*, 29, 85-95.
- Russell, J. A. (1980), A Circumplex Model of Affect. *Journal of Personality and Social Psychology*, 39, 6, 1161 - 1178.

- Slovic, P. & Lichtenstein, S. (1971), Comparison of Bayesian and regression approaches to the study of information processing in judgment, *Organizational Behaviour & Human Performance*, 6, 649-744.
- Tractinsky, N., Katz, A. & Ikar, D. (2000), What is Beautiful is Usable, *Interacting with Computers*, 13, p.127-145.
- Zajonc, R. (1980), Feeling and Thinking: Preferences need no References, *American Psychologist*, February, p.151.