Why do I like it? The relationships between icon characteristics, user performance and aesthetic appeal

Siné J.P. McDougall
Psychology, School of Design Engineering & Computing
Bournemouth University. UK

Irene Reppa
Department of Psychology
Swansea University, UK

Until recently the guiding tenet in human-computer interaction was that any interface must be easy to learn and use. However, it has been increasingly recognized that the appeal of the interface to the user and their enjoyment of it is also important. The aim of the current study was to examine the nature of the relationships between icon characteristics, user performance, and aesthetic appeal. When participants were asked to rate the appeal of a corpus of icons, it was found that the same icon characteristics predicted appeal as those predicting user performance. The theoretical and practical implications of the remarkable similarity in the factors determining appeal and usability are discussed.

One of the fundamental aims of any human-computer interaction system is to facilitate fast and efficient performance. Particular emphasis has been given to interface usability and performance measures during tasks such as time to learn, levels of accuracy, and response latencies (e.g., Butler, 1996). However, there is a growing recognition that enhancing the aesthetic appeal of an interface may be just as important as improving its usability (e.g., Hassenzahl & Tractinsky, 2006). Recent research has shown that performance and appeal are inter-related (e.g., Kurosu & Kashimura, 1995; Lingaard & Dudek, 2003; Tractinsky, Katz, & Ikar, 2000; Tractinsky, 2004; Wiedenbeck, 1999). For example, our pleasure in using a system can determine how much effort we are likely to put in when learning it and increases user performance for interfaces we already know (e.g., Wiedenbeck, 1999; Tractinsky, Katz, & Ikar, 2000). Conversely, research has shown that aesthetically pleasing designs tend to increase our perceptions of system usability (e.g., Kurosu & Kashimura, 1995; Tractinsky, 2004).

Although the general relationship between usability and aesthetics is well documented, the interface characteristics that that can affect aesthetic appeal, and their relationship with user performance, have not been examined. This is important if both interface usability and appeal are going to be enhanced. This study therefore examined the relationship between icon characteristics that are known to influence user performance and their effect on aesthetic appeal. Icons were used because (a) they are employed to communicate information on a wide variety of interfaces and (b) the effects of icon characteristics on performance are already well known (e.g. Isherwood & McDougall, 2007; McDougall et al, 2000). Throughout this paper the term aesthetic *appeal* is used, rather than aesthetic *preference*, since the former refers to the power to attract or arouse interest, while preference refers to selecting one thing over another. In the study reported here participants were asked to rate appeal, rather than choose between items.

ICON CHARACTERISTICS & PERFORMANCE

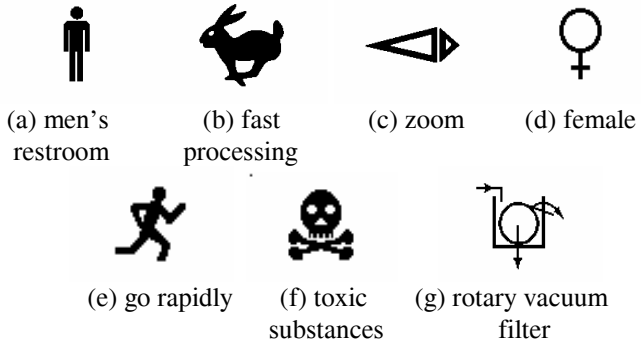
The visual complexity, concreteness, and familiarity of icons have all been shown to affect user performance.

Visual complexity

Simple icons enhance performance because they can be discriminated more easily in arrays (e.g., Byrne, 1993) and are located more easily in visual search (Byrne, 1993; Scott, 1993). Research also suggests that although visual complexity has an important role to play in search it is not directly involved in icon identification (McDougall et al., 2000).

Concreteness

When icon identification is the key element of a task, concreteness (i.e. pictorialness) appears to be an important determinant of the speed and accuracy with which users can identify icons (e.g. Green & Barnard, 1990; Rogers & Oborne, 1987; Stotts, 1998). This is because concrete icons depict objects, allowing people to use their knowledge of the everyday world in order to interpret them but this is less easily done with abstract icons (c.f. Figure 1a and b with c and d). Until recently concreteness was sometimes seen as an icon's most important property, however research now suggests that the effects of concreteness on user performance are less than previously thought (Isherwood & McDougall, 2007). This is because only a limited number of functions can be represented concretely and getting a close fit between pictures and functions is not always easy. For example, naming the rabbit in Figure 1b does not allow you to arrive at its intended meaning.



(a) men's restroom (b) fast processing (c) zoom (d) female

(e) go rapidly (f) toxic substances (g) rotary vacuum filter

Figure 1: Examples of icons

Familiarity

Apart from icon complexity and concreteness, another important determinant of performance in icon tasks is the user's familiarity with what is depicted in the icon. Isherwood & McDougall (2007) found that familiarity was an important predictor of speed and accuracy of icon identification, irrespective of concreteness. For example, our familiarity with the icon representing 'female' in Figure 1d allows us to identify it more quickly and effectively compared with the icon representing 'fast processing' (Figure 1b). Furthermore, Forsythe, Mulhern & Sawey (2008) have shown that there is a correlation between icon familiarity and visual complexity: familiar icons are perceived as being simpler.

Figure 2 illustrates the findings of research to date examining the effects of visual complexity, concreteness and familiarity on user performance. Familiarity is an important determinant of user performance and research suggests that it encompasses the effects previously attributed to concreteness. Visual complexity also determines user performance via its role in visual search and is correlated with, but not encompassed by, icon familiarity.

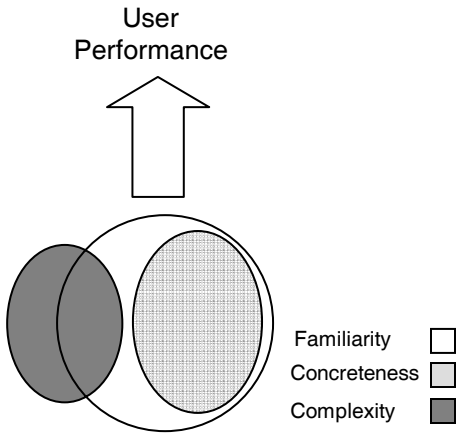


Figure 2: Relationships between icon characteristics and performance

ICON CHARACTERISTICS & AESTHETIC APPEAL

There are remarkable parallels between the stimulus characteristics which affect performance and those which determine aesthetic appeal. These are as follows:-

(i) Stimulus complexity has a significant influence on aesthetic appeal judgments (e.g., Berlyne, 1974; Jacobsen & Hofel, 2002).
(ii) Concrete, or representational, pictures are also known to be preferred to abstract ones (e.g., Vartanian & Goel, 2004; Kawabata & Zeki, 2004).
(iii) Familiarity with the stimulus can also influence its aesthetic appeal to the observer. Zajonc (1968) was the first to show that even when stimuli were presented only once for a brief period, the appeal for those stimuli was increased relative to others that had not been previously presented. The difference in liking between old and new stimuli has been termed the *mere exposure effect (*Zajonc, 1968) and refers to the effect of familiarity (implicit or explicit) on appeal.

What is not known is whether or not these findings, which suggest the possibility of specific relationships between icon characteristics, user performance and appeal, generalize and apply to icons. Logically, *if appeal follows performance*, we should expect that predictors of aesthetic appeal should be similar to those observed for user performance in Figure 2. If this were the case we would expect that (a) the variance in appeal explained by familiarity would encapsulate that predicted by icon concreteness and (b) there would be an overlap in the variance predicted by visual complexity and familiarity. This complex pattern of findings could not be predicted on the basis of previous aesthetics research but follows logically from our premise that performance and appeal are inter-related.

METHOD

Participants

A total of forty participants, who were undergraduate and postgraduate students at the University of Wales Swansea, volunteered to take part in this study. Six participants did not follow instructions to use the full range of the 1-5 rating scale (providing the same rating for over 80% of the icons). Their data was therefore excluded from subsequent analyses. The mean age of the remaining participants was 23.4 years (SD=3.5 years; 28 females and 6 males).

Materials

The set of 239 icons and symbols used by McDougall, Curry & de Bruijn (1999) were used because ratings had already been obtained for a number of characteristics:

(i) Concreteness: the extent to which icons depicted objects, materials or people
(ii) Visual complexity: the amount of detail or intricacy in the icon
(iii) Familiarity: the extent to which icons were perceived as familiar
(iv) Meaningfulness: the relationship between what is depicted in the icon and the function it refers to (the closer the relationship the more meaningful the icon is).

The original corpus of icons was chosen from a wide variety of sources to ensure that they were representative of the broad spectrum of applications in which icons and symbols are used. They were also selected to ensure a wide distribution across each of the above characteristics (concrete and abstract icons, complex and simple icons, etc).

Procedure

Participants were asked to rate the icons on a 1-5 scale in accordance with how much they liked them (1=really dislike,

5=really like). Unless their response was neutral (3=neither like nor dislike), participants were also asked to indicate what it was about the icon that made it either appealing or unappealing. Participants were instructed to use the full range of ratings from 1-5. Because of the large number of icons involved, participants were divided into two equal groups and asked to rate half of the corpus. Two booklets were created, one for each half of the corpus, and the order in which participants were assigned booklets was counterbalanced. Icons were presented in 25-page booklets with five icons on each page. Icons were presented in random order on each page. Pages were assembled into booklets in accordance with a Latin square design to ensure that each participant was presented with the icons in a different order.

RESULTS

We analysed the data by-items, rather than by-subjects, because it was the nature of the icons which were the focus of interest. The mean aesthetic rating was 3.01 (SD=0.47, Min=1.88, Max=4.53) and ratings were normally distributed. Few participants reported really disliking icons (a rating of 1). This was perhaps not surprising given the effort that usually goes into their design.

The extent to which participants reported liking icons was then correlated with ratings of icon characteristics previously obtained by McDougall et al. (2000). All correlations were significant: r(concreteness)=.32, p<.01; r(visual complexity)= -.29, p<.01; r(familiarity)=.46, p<.01. These correlations suggest that concreteness, visual complexity and familiarity all contribute to determining aesthetic appeal and that individuals favour icons which are represent familiar, pictorial, items in a simple way.

In order to examine the extent to which each characteristic had an independent, or unique, role in determining aesthetic appeal we carried out a series of regression analyses with aesthetic ratings as the dependent variable, in which each characteristic was entered in turn as the final variable in the regressions. Table 1 summarises the findings of these analyses.

In the first regression ratings of icon concreteness were entered first into the regression, followed by visual complexity and familiarity. In this analysis concreteness appeared to be an important determinant of aesthetic appeal accounting for over 10% of the variance seen in aesthetic ratings. However, in the second analysis where familiarity was entered first into the regression, concreteness explained little of the variance in appeal once familiarity with the icon was taken into account. Visual complexity appeared to have a smaller, but statistically reliable role, in determining appeal which was independent, to some extent at least, of icon familiarity. Taken together these findings suggest that familiarity is of primary importance in determining appeal but that keeping icons simple (since complexity is inversely related to appeal) is also important. The nature of the overlapping inter-relationships between icon characteristics in determining appeal is therefore very similar to that depicted in Figure 2 except that in this instance it is aesthetic appeal, rather than performance, which is being predicted.

Steps	Variable	% Variance
1	Concreteness	10.3**
2	Visual complexity	7.6**
3	Familiarity	5.3**
1	Familiarity	21.0**
2	Concreteness	0.2
3	Visual complexity	2.0*
1	Visual complexity	8.7**
2	Familiarity	14.5*
3	Concreteness	0

**p<.01; *p<.05

Table 1: Summary of Fixed-Order Stepwise Regression Analyses

This pattern of results was also reflected in the next set of analyses where binomial tests and chi-square analyses were carried out on the frequency with which individuals reported making aesthetic judgements about icons (see Table 2).

Reasons for aesthetic judgements	% Times Reported
Known/familiar object in icon	15.77
Meaning associated with known object	18.20
No object/shape could be identified	2.38
Complexity/simplicity	9.51
Global features (overall pattern, symmetry)	14.45
Local features (individual features identified)	14.31
Unclassified other	25.38

Table 2: Frequency with which different reasons for aesthetic judgements were reported

When individuals were familiar with the object depicted in the icon, they were much more likely to report finding it attractive (p<.001). For example, an icon depicting a fan was identified as either a fan or a flower (the fan blades were seen as petals) and rated positively. On the few occasions when participants reported being *unable* to identify an object or shape in the icon or the icon as a reason for making an aesthetic judgement all, without exception, found this unattractive.

Participants often reported that the meaning associated with objects they identified in the icon led to an aesthetic judgement. This judgement was equally likely to be a positive or a negative judgement (p>.05). For example, the icon shown in Figure 1e which often appears in exit signs, was often seen as being associated with sport and, therefore, had positive connotations while Figure 1f appeared threatening and had negative connotations.

Participants were much more likely to find simple icons appealing but found complexity or intricacy in an icon

unappealing ($\chi^2(1)=81.73$, $p<.001$; see, for example, the complex icon in Figure 1g).

Although participants reported using global or local features within the icons to make aesthetic judgements equally frequently, they were more likely to report finding the overall pattern, shape, or symmetry of icons attractive and were more likely to report individual features as unattractive ($\chi^2(1)=63.16$, $p<.001$).

Participants were equally likely to report reasons for liking or disliking an icon ($p>.05$).

DISCUSSION

This study examined the characteristics that are associated with the aesthetic appeal of icons. Given the remarkable correspondences between the stimulus characteristics affecting both performance and appeal, it was hypothesized icon characteristics which affected user performance would have similar effects on aesthetic appeal. This was indeed what was found. First, icon familiarity accounted for most of the variance in appeal. Second, icon concreteness did not predict appeal once the effects of icon familiarity were accounted for. Third, visual complexity accounted for a small but significant amount of the variance although this overlapped to some degree with icon familiarity.

Reports from individuals about the reasons for their appeal judgments also mirrored these findings. However, other dimensions also emerged from the frequency data that inform us about the nature of representations that underlie aesthetic appeal. Participants were more likely to explain the appeal of a stimulus in terms of its global shape (e.g., symmetry) and lack of appeal in terms of local features (e.g., an arrow at the top of the icon, or a dark dot in the icon). How does this relate to familiarity? Familiar stimuli tend to be processed in a more global (or configural) manner than unfamiliar stimuli (e.g., Farah, Wilson, Drain & Tanaka, 1998). Once a stimulus is processed globally then it is difficult to break it down to its individual components. On the other hand, unfamiliar stimuli are more likely to be processed in terms of their individual components. Therefore, it is possible that there is an implicit relationship between appeal for familiar stimuli and the type of visual processing of the stimulus (global or local). Clearly this possibility requires further investigation.

Theoretical Accounts

Our findings cannot be easily accounted for in terms of existing theoretical accounts, which tend to explain user performance or appeal rather than both. For example, Isherwood & McDougall (2007) used Johnson, Paivio & Clark's (1996) theory of picture naming to explain the speed and accuracy with which icons were identified. Visual complexity was associated with the first processing stage - the search and location of an item – while familiarity was linked to the ease with which matching visual representations could be accessed from long term memory. Johnson et al's theory, however, is silent with respect to aesthetic appeal.

A more promising possibility is the *perceptual fluency hypothesis*, which is a commonly accepted account of the mere exposure effect. It postulates that visual stimuli which have been previously encountered are more quickly and effectively processed perceptually, leading to the subjective experience of processing fluency (Reber et al, 1998). If individuals are asked to evaluate a stimulus, then they are more likely to attribute this fluency to liking and provide higher liking scores for such stimuli (Bornstein & D'Agostino, 1994; Seamon, Brody & Kauf, 1983). Simple, concrete and familiar icons are all processed more quickly than complex, abstract and unfamiliar icons. This could explain the ratings of aesthetic appeal which were observed. However, it is more difficult to explain the complex pattern of findings observed in performance and appeal (as in Figure 2).

We propose, instead, that our results can be accounted for by the *processing fluency hypothesis*. This hypothesis recognizes the basic premise that certain icon characteristics can influence aesthetic appeal evaluations, but proposes that these evaluations may ultimately depend on user performance. A key difference between this account and the *perceptual fluency* account is that it is a dynamic, rather than static, explanation of aesthetic appeal. Implicit within the perceptual fluency account is the assumption that stimuli elicit a single aesthetic value rather than one which might change depending on the task or context. This assumption arises, at least to some extent, because studies examining perceptual fluency have presented items singly rather than in an array as they might appear on an interface. The processing fluency hypothesis therefore departs theoretically from processing fluency account because it assumes that user performance, and subsequently aesthetic appeal, will be influenced by the task demands and the context, or interface array, in which an icon appears. This, however, needs to be tested explicitly in future research.

Practical Implications and Future Research

These findings challenge both researchers and interface designers to take account of the strong links observed between user performance and appeal. Indeed, in doing so it might be tempting to assume that if the factors affecting user performance are known, then the factors affecting appeal are also known. This, however, would not take account of the following:-

(1) That other, additional factors, may affect aesthetic appeal. This was apparent in our examination of the reasons for aesthetic appeal (see Table 2).

(2) Where usability is paramount, appeal may not be so closely correlated with performance (e.g. in air traffic control displays task demands may determine the closeness of the relationship between usability and appeal).

(3) Research has shown that the factors affecting icon usability change as the icon set is learned. Whether such changes occur in appeal is not yet known.

In conclusion, our results show that icon characteristics which influence user performance, such as complexity, concreteness and familiarity, can also influence aesthetic appeal and that they do so in complexly similar ways which are difficult to predict on a purely intuitive basis. On the basis of these results we suggest that, in order to optimize interface design, the current focus on enhancing interface appeal needs to be closely tied to a sound understanding of the factors influencing user performance. We propose that the processing fluency hypothesis has the potential to provide a theoretical framework that makes testable predictions regarding this dynamic relationship between performance and aesthetic appeal.

REFERENCES

Berlyne, D.E. (1974). *Studies in the new experimental aesthetics*. Washington, DC: Hemisphere.

Bornstein, R.F. & D'Agostino, P.R. (1994). The attribution and discounting of perceptual fluency: Preliminary tests of a perceptual fluency/attributional model of the mere exposure effect. *Social Cognition, 12*, 103-128.

Butler, K.A. (1996). Usability engineering turns 10. *Interactions, 3*, 59-75.

Byrne, M. D. (1993). Using icons to find documents: Simplicity is critical. *Proceedings of INTERCHI '93*, 446 – 453.

Farah, M. J., Wilson, K. D., Drain, M. & Tanaka, J. N. (1998). What is "special" about face perception? *Psychological Review. 105*, 482-498

Forsythe, A., Mulhern, G. & Sawey, M. (2008). Confounds in pictorial sets: the role of complexity and familiarity in basic-level picture processing. *Behavior Research Methods*, *40*, 116-129.

Green, A. J. K. & Barnard, P. J. (1990). Iconic interfacing: The role of icon distinctiveness and fixed or variable screen locations. In D. Diaper et al. (Eds.), *Human-computer interaction – Interact '90*. North-Holland: Elsevier.

Hassenzahl, M. & Tractinsky, N. (2006). User experience: A research agenda. *Behavior & Information Technology*, *25***,** 91-97.

Isherwood, S., McDougall, S. & Curry, M. (2007). Icon identification in context: The changing role of icon characteristics with user experience. *Human Factors, 49*, 465-476.

Jacobsen, T. & Hofel, L. (2002). Aesthetic judgments of novel graphic patterns: Analyses of individual judgments. *Perceptual & Motor Skills, 95*, 755-766.

Johnson, C.J., Paivio, A., & Clark, J.M. (1996). Cognitive components of picture naming. *Psychological Bulletin, 120*, 113-139.

Kawabata, H. & Zeki, S. (2004). Neural correlates of beauty. *Journal of Neurophysiology, 91*, 1699-1705.

Kurosu, M. & Kashimura, K. (1995). Apparent usability vs. inherent usability. *CHI'95 Conference Companion*, 292-293.

Lindgaard, G. & Dudek, C. (2003). What is this evasive beast we call user satisfaction? Interacting with Computers, 15, 429-452.

McDougall, S., Curry, M.B. & de Bruijn, O. (1999). Measuring symbol and icon characteristics: Norms for concreteness, complexity, meaningfulness, familiarity and semantic distance for 239 symbols. *Behavior Research Methods, Instruments & Computers*, *31*, 487-519.

McDougall, S., de Bruijn, O. & Curry, M. (2000). Exploring the effects of icon characteristics on user performance: The role of concreteness, complexity and distinctiveness. *Journal of Experimental Psychology: Applied*, *6*, 291-306.

Reber, R., Winkleman, P. & Schwarz, N. (1998). Effects of Perceptual Fluency on Affective Judgments. *Psychological Science, 9*, 45 – 48.

Rogers, Y. & Oborne, D.J. (1987). Pictorial communication of abstract verbs in related to human-computer interaction. *British Journal of Psychology, 78,* 99-112.

Scott, D. (1993). Visual search in modern human-computer interfaces. *Behaviour and Information Technology, 12*, 174 – 189.

Seamon, J.G., Brody, N. & Kauf, D.M. (1983). Affective discrimination of stimuli that are not recognized: Effects of shadowing, masking, and cerebral laterality. *Journal of Experimental Psychology: Learning, Memory & Cognition, 12*, 544-555.

Stotts, D.B.(1998). The usefulness of icons on the computer interface: Effect of graphical abstraction and functional representation on experienced and novice users. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, 453-457. Santa Monica: Human Factors Society.

Tractinsky, N., Katz, A.S. & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers, 13*, 127-145.

Tractinsky, N. (2004). A few notes on the study of beauty in HCI. *Human-Computer Interaction*, *19*, 351-357.

Vartanian, O. & Goel, V. (2004). Neuroanatomical correlates of aesthetic preferences for paintings. *Neuroreport, 15*, 893-897.

Wiedenbeck, S. (1999). The use of icons and labels in an end user application program: an empirical study of learning and retention. *Behaviour Information Technology, 18*, 68-82

Zajonc, R.B. (1968). Attitudinal effects of mere exposure. *Journal of Personality & Social Psychology Monographs, 9*, 1-27.