

What makes icons appealing? The role of processing fluency in predicting icon appeal in different task contexts

Siné McDougall^{a1}, Irene Reppa^{b1}, Jozef Kulik^a, Alisdair Taylor^{a2}

^aPsychology Department, Faculty of Science & Technology, Bournemouth University, Fern Barrow, Poole BH12 5BB, UK.

Tel: +44 1202 961722 Email: smcdougall@bournemouth.ac.uk

^bPsychology Department, Swansea University, Swansea, SA2 8PP, UK.

Tel: +44 1792 295963 Email: i.reppa@swansea.ac.uk

Abstract

Although icons appear on almost all interfaces, there is a paucity of research examining the determinants of icon appeal. The experiments reported here examined the icon characteristics determining appeal and the extent to which processing fluency – the subjective ease with which individuals process information – was used as a heuristic to guide appeal evaluations. Participants searched for, and identified, icons in displays. The initial appeal of icons was held constant while ease of processing was manipulated by systematically varying the complexity and familiarity of the icons presented and the type of task participants were asked to carry out. Processing fluency reliably influenced users' appeal ratings and appeared to be based on users' unconscious awareness of the ease with which they carried out experimental tasks.

Keywords: icons; processing fluency; user experience; aesthetics; appeal

¹ Corresponding authors

² Now at the University of British Columbia

1 Introduction

Over the last 15-20 years there has been a shift in focus in human-computer interaction research, away from the sole emphasis on usability and performance towards a recognition that users' affective experience of interfaces can be at least as, and sometimes more, important than usability (e.g. Hartmann et al., 2008; Hassenzahl & Tractinsky, 2006; Mack & Sharples, 2009; Schmidt et al., 2009; Tractinsky et al., 2000). This has been particularly true of research examining website design where attracting consumers to appealing websites is paramount (Golander et al., 2012). A great deal of research has focused on getting to grips with the dizzying array of website characteristics that may contribute to user perceptions of website appeal (e.g. Cyr et al., 2010; De Wulf et al., 2006; Lavie & Tractinsky, 2004; Moshagen & Theilsch, 2010; Thuring & Mahlke, 2007; van Schaik & Ling, 2005). In contrast, there has been little research which has investigated determinants of icon appeal³ despite the fact that they are an integral part of virtually every interface. The aim of the experiments reported here was to examine the extent to which processing fluency might provide an explanation of users' perceptions of appeal.

1.1 Rapid appeal evaluations

It is now well established that users can make very rapid - almost instantaneous - evaluations of the appeal of interfaces (Handy et al., 2008; Lindgaard et al., 2006; Lindgaard et al., 2011). Lindgaard et al. (2006) interpreted participants' ability to make these rapid judgements as being the result of the 'mere exposure effect'. Zajonc (1968) first showed that merely seeing something, even for the briefest of moments, is sufficient for stimulus preferences to develop. These instantaneous affective responses are therefore not thought to be the result of deliberate thought or reflection but occur automatically without conscious awareness.

In an event-related potential (ERP) study, Handy et al. (2008) examined whether processing of appeal occurred automatically with icons and logos. Participants in Handy et al.'s study were shown logos and asked to press a key when a given target logo appeared among a series of non-target logos appeared. During this task ERP responses were noted and it was only *after* the EEG recording equipment was removed that participants were asked to indicate their liking of the *non-target logos* using a 1-7 Likert scale. The key finding in this study was that ERP responses to logos within the first 200ms of stimulus onset varied as a function of the appeal ratings which were obtained later. Handy et al. argued that we 'rapidly and implicitly evaluate ... images at a hedonic level' (p. 124).

1.2 Processing fluency and appeal evaluations

As users move beyond these first impressions, one theoretical approach which might explain changes in appeal over time is processing fluency theory (Alter & Oppenheimer, 2009; Labroo Dhar & Schwarz, 2008; Oppenheimer, 2008; Reber et al., 2004). Processing fluency refers to our experience of ease or difficulty when carrying out mental tasks. If an object is very small and distant or simply out of focus, we are aware it is hard to see; if an event is

³ Throughout this paper the term appeal refers to mild aesthetic experiences made on the basis of simple judgements of liking by participants (see Reber et al., 2004 for a review). Liking and appeal refer to the power to attract whereas preference refers to selecting one thing over another. While similar mechanisms are thought to underpin both liking judgements and preferences, our focus in this paper is on judgements of liking/appeal.

easy to retrieve from memory, we have an implicit awareness of how easy it was to access. Because this feeling of relative ease or difficulty in processing is generated by almost any cognitive process and is almost effortless to access, it can act as a cue in making judgements in a wide variety of situations (Alter & Oppenheimer, 2009; Oppenheimer, 2008).

There is compelling evidence that processing fluency when dealing with visual stimuli can act as a cue in determining judgements of appeal and preferences across a wide range of tasks (e.g. Labroo et al., 2007; Reber et al., 2004). For example, when individuals were shown stimuli against a highly contrastive background they found them more appealing than identical stimuli against a less contrastive background and pictures presented for 400ms were preferred in comparison to those shown for 100ms (Reber et al., 1998). When Constable et al., (2013) examined preferences for hand-painted mugs, they found that the speed with which individuals were able to find a particular mug within a search array determined their later judgements of appeal and therefore concluded that processing fluency plays a role in forming preference judgments.

As we learn stimuli, and find them easier to process and access from memory, our liking for them increases. Zizek and Reber (2004) asked participants to evaluate the appeal of letter strings after an implicit grammatical learning task. In the evaluation phase, when participants were asked to rate their liking of the letter strings, those which followed the same grammatical rules – even if not shown previously – were evaluated more positively than those which did not follow the grammatical rules participants had learned. This could not easily be attributed to mere exposure since the strings following similar grammatical rules had not been seen before. Zizek and Reber therefore attributed these findings to the greater processing fluency resulting from their understanding of the grammatical rules that participants had learned. Similar findings have been reported in other studies of learned ‘grammars’ (Gordon & Holyoak, 1983; Manza & Bornstein, 1995; Manza et al., 1998).

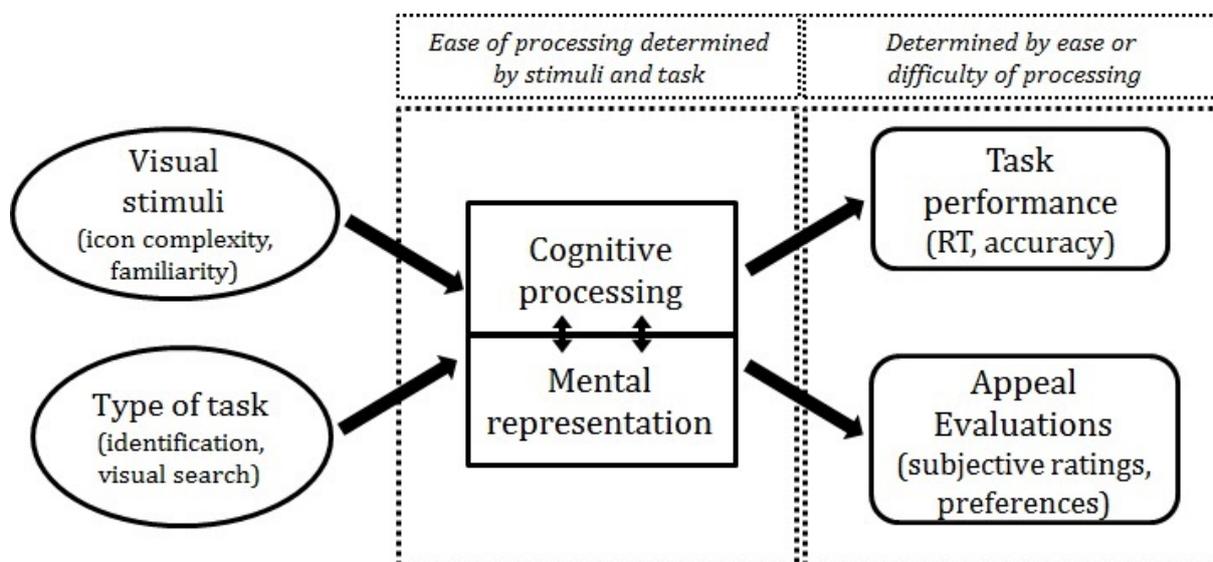


Figure 1: The ease, or fluency, with which visual stimuli are processed cognitively is determined by the characteristics of the stimuli and the tasks users are carrying out. Ease of processing determines task performance and appeal evaluations, hence evaluations of appeal reflect task performance and vice versa.

1.3 Experimental rationale

Taken together previous findings suggest that, if we find icons on an interface easy or difficult to process, then this will be reflected in the judgements we make about how appealing they are, i.e. processing fluency acts as an implicit heuristic to determine appeal judgements (i.e. preferences, subjective ratings of appeal or liking). Since patterns of performance measured using response times and accuracy *also* depend on ease of processing (see Figure 1), it follows that appeal judgements should reflect patterns of performance made to those stimuli. More specifically, ratings of liking for icons on interfaces should reflect the speed and accuracy with which we can search for and identify icons because they both derive from ease of processing. The experiments reported here were designed to examine whether or not this was the case and the extent to which processing fluency can explain judgements of icon appeal. In the experiments which follow participants were asked to provide subjective ratings of ‘liking’ for icons rather than aesthetic appeal. ‘Liking’ was thought to be more appropriate since this is a more everyday term for appeal evaluations and more commonly used (see Constable et al., 2013, for a similar approach).

The first step in our investigation was to examine whether or not there was any kind of relationship between appeal and performance for icons. To that end, Experiment 1 sought to identify the icon characteristics predicting appeal and whether or not they are the same as known to those predict user performance. Experiments 2-4 examined whether or not judgements of appeal would reflect the changes in performance resulting from the nature of the icon stimuli presented and the tasks being carried out. Two tasks were designed to mimic searching for icons on an interface. Experiment 2 used a *search* task in which participants were presented with an icon and required to find the matching icon in an array, mimicking the search for icons with particular functions on displays. Experiments 3 and 4 used an *identification* task in which participants were given the function, or name, of the icon and then asked to find the matching icon in an array. The identification task therefore involved learning the meaning of the icons over a series of experimental trials by associating the function with the appropriate icon, reflecting the learning process that users go through as they gain experience with icon sets on interfaces. Given that the icon stimuli were the same for the search and identification tasks, it was expected that changes in evaluations of appeal between Experiment 2 versus Experiments 3 and 4 would be the product of differences in ease of processing resulting from the change in task. In Experiment 4 the icon stimuli and experimental task were the same as for Experiment 3, however, subjective ratings of liking were obtained twice from each participant, early and later in experimental trials. The aim of this manipulation was to give greater emphasis to participants’ growing experience with the icons and the task over experimental trials.

1.4 Icon characteristics and task performance

Icon characteristics that have previously been shown to affect user performance include the visual complexity, concreteness, and familiarity of the icons.

1.4.1 Visual complexity

Research has shown that simple icons are found faster on an interface array than complex ones (e.g. Byrne, 1989; McDougall et al., 2000; McDougall & Isherwood, 2009; McDougall et al., 2006, Scott, 2003). Explanations of differences in search time between

simple and complex icons often rely on Treisman’s feature integration theory (Treisman & Gelade, 1980; Treisman & Gormican, 1988). In this view, there are two stages in visual search; the first involves fast parallel processing of primitive visual features in a scene, whereas the second stage involves slower serial processing that binds primitive features together into an object. From this, it follows that the more visually complex an icon is (and therefore the more features it has), the more time will be involved in binding features together, producing slower response times for more complex signs. However, Forsythe et al. (2008) have shown that there is a correlation between icon familiarity and visual complexity and that this is the result of familiar icons being perceived as simpler (see also Bates et al., 2003; Chi & Dewi, 2014). This suggests that searching for icons in arrays may have a top-down, as well as a bottom-up, processing component and aligns with later theoretical approaches to search which have emphasised the importance of prior knowledge of shapes, objects, or complex scenes (Bundesen, 1998; Green & Oliva, 2009; Joubert et al., 2007; Logan, 1996; Malcolm & Henderson, 2009; Rousselet et al., 2004; Wolfe et al., 2011). Despite this correlation, both icon familiarity and complexity *independently* predict performance when participants are asked to find icons in an array (Isherwood & McDougall, 2007; McDougall & Isherwood, 2009). Thus, visual complexity appears to be important in determining search time for icons on interfaces and is related to familiarity because search is driven, in part at least, by our prior knowledge of stimuli.

1.4.2 Concreteness

Both the accuracy and speed with which users can identify icons is affected by the concreteness, or pictorialness, of the icons on an interface (e.g. Chan & Ng, 2010; Green & Barnard, 1990; Leung et al., 2011; Rogers & Osborne, 1987; Schroder & Zeifler, 2008; Stotts, 1998). This is thought to be because concrete icons depict real world objects, allowing people to use their prior knowledge to discern their meaning (c.f. Figure 2a and b with c and d).

While research suggests that concreteness is the most important determinant of performance, other work has shown that the effects of concreteness on user performance are less than previously thought (Isherwood & McDougall, 2007; McDougall & Isherwood, 2009). This appears to be because only a limited number of functions can be represented concretely - getting a close fit between pictures and functions is not always easy. For example, naming the rabbit in Figure 2b does not make it possible to arrive at its intended meaning, the association between the icon and its intended meaning needs to be learned.

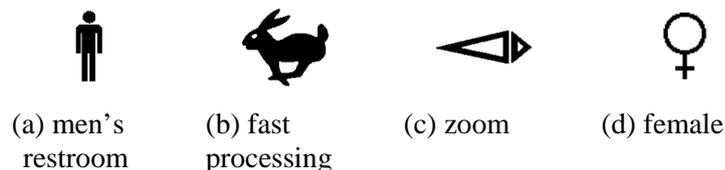


Figure 2: Examples of different types of icons

1.4.3 Familiarity

Research has consistently shown that icon familiarity determines the speed and accuracy with which icons and objects can be identified (Ben-Bassat & Shinar, 2006; Chan & Ng, 2010; Lesch et al., 2011; Lui, 2005; Shinar et al., 2003; Wang, Cavanagh & Green, 1994; Wolfe & Alvarez, 2011). When compared to other icon characteristics, it appears to be the

most important determinant of ease of identification (Isherwood & McDougall, 2007; McDougall & Isherwood, 2009). For example, our familiarity with the abstract icon representing ‘female’ in Figure 2d allows us to identify it more quickly and effectively compared with the pictorial rabbit icon representing ‘fast processing’ (Figure 2b). In the picture naming literature, rated familiarity of the pictures is also a strong and consistent predictor of naming times and is thought to reflect the ease with which semantic representations can be accessed from long-term memory (Alario et al., 2004; Bates et al., 2003; Ellis & Morrison, 1998; Lambon-Ralph et al., 1998; Snodgrass & Yuditsky, 1996). Thus, when icons are presented, ease of processing is likely to be determined by the extent to which interface users are able to access an appropriate function, or meaning, reflecting our familiarity with the icon-function relationship.

Figure 3 summarises the findings of research to date examining the effects of visual complexity, concreteness and familiarity on user performance. Familiarity with the icon is the most important determinant of user performance and research suggests that it encompasses the effects previously attributed to concreteness. This is because the function of an icon cannot always be inferred on the basis of a concrete pictorial and there are many functions represented by icons where there is no easy fit between pictures and functions. In such circumstances, familiarity is the key determinant of ease of processing. Visual complexity is also important in determining the ease with which we are able to search for icons on an interface. Visual complexity and familiarity are correlated to some extent reflecting the degree to which search is driven by our prior knowledge of stimuli.

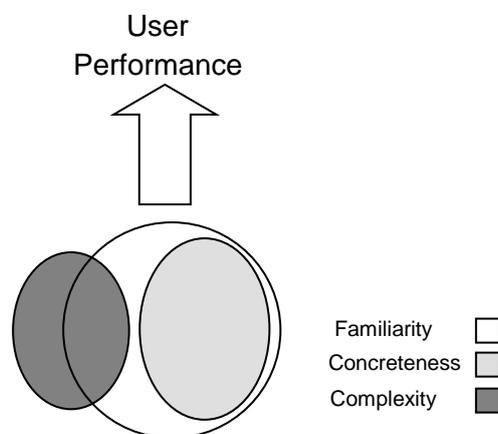


Figure 3: Relationships between icon characteristics and performance

1.5 Stimulus characteristics that might affect icon appeal evaluations

Although very little is known about the stimulus characteristics affecting icon appeal, concreteness, complexity and familiarity are known to be important determinants of our aesthetic and appeal responses to other types of stimuli, such as works of art or websites. Concrete, or representational, pictures are preferred to abstract ones (e.g. Kawabata & Zeki, 2004; Vartanian & Goel, 2004); the visual complexity of pictures, websites and graphic patterns has a significant influence on aesthetic appeal judgements (e.g. Berlyne, 1974; Bauerly & Liu, 2008; Jacobsen & Hofel, 2002; Lai et al., 2010; Moshagen & Theilsch, 2010, 2012; Orth & Wirtz, 2014); and, of course, familiarity enhances appeal for a wide range of

stimuli although this depends to some extent on whether or not the items were liked initially (Bornstein, 1989; Fang et al., 2007; McCoy et al., 2009; Monahan et al., 2001; Zajonc, 1968, 2001).

2 Experiment 1: Icon characteristics determining appeal

The aim of this experiment was to examine the extent to which icon characteristics known to predict user performance would also predict evaluations of appeal. If appeal judgements derive from processing fluency heuristics then predictors of appeal should be similar to those observed for user performance in Figure 3. On this basis we would expect that (a) icon familiarity would be the strongest predictor of appeal (b) icon concreteness would be a less important predictor of appeal than familiarity (c) icon complexity would also predict appeal judgements and (d) there would be an overlap in the variance predicted by visual complexity and familiarity. This pattern of findings follows from our current understanding of icon characteristics determining user performance.

2.1 Method

2.1.1 Participants

Forty Swansea University undergraduate and postgraduate volunteers took part in this study. Six participants' data was excluded because they did not follow instructions to use the full range of the 1-5 rating scale and provided the same rating for over 80% of the icons. The mean age of the remaining participants was 23.4 years ($SD=3.5$ years; 28 females and 6 males).

2.1.2 Materials

A set of 239 icons and symbols used by McDougall, Curry & de Bruijn (1999) were used because subjective ratings had already been obtained for a number of characteristics including the perceived (i) visual complexity (ii) concreteness and (iii) familiarity of the icons. 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. These included public information signs, symbols for use on electrical equipment, websites, computer interfaces, vehicle and aircraft controls and displays, and military symbols. Icons in the corpus were also selected to ensure a wide distribution across each of the icon characteristics under consideration (i.e. complex vs simple icons, concrete vs abstract icons, familiar vs unfamiliar icons). Ratings of each characteristic were obtained using 5-point Likert scales: for visual complexity 1=very simple, 5=very complex; for concreteness 1=definitely abstract, 5=definitely concrete; and for familiarity 1=very unfamiliar, 5=very familiar).

2.1.3 Procedure

In this experiment, 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. As noted earlier,

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 icons presented in random order on each page and 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.

2.2 Results and Discussion

The data was analyzed using by-items, rather than by-subjects, analyses because this type of analysis allowed us to examine icon ratings across the full corpus of icons and it was appeal for icons in the corpus which formed the focus of this study. Ratings of liking were distributed normally and the mean rating for all 239 icons was 3.01 (SD=0.47, Min=1.88, Max=4.53).

2.2.1 Correlations between icon characteristics and appeal

The extent to which participants reported liking icons was correlated with ratings previously obtained of icon complexity, concreteness and familiarity. All correlations were significant: $r(\text{visual complexity})=-.29, p<.01$; $r(\text{concreteness})=.32, p<.001$; $r(\text{familiarity})=.46, p<.001$.

2.2.2 Regression analyses

Table 1: Summary of Fixed-Order Stepwise Regression Analyses

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

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

In order to investigate the inter-relationships between these variables further, a series of regression analyses were carried out. The aim of these analyses was to examine the extent to which each characteristic had an independent, or unique, role in determining aesthetic appeal. A series of regression analyses with ratings of liking as the dependent variable, in which each characteristic was entered in turn as the final variable in the regressions (see Table 1). Those variables which remained significant predictors of appeal, even after the effects of other variables had been taken into account could be taken as having a unique role in determining appeal.

In the first regression, when icon concreteness was entered first into the regression, it was a significant predictor of appeal. 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 appeal, rather than performance, which is being predicted. This pattern of results was also reflected in the reasons participants gave for either liking or disliking icons.

Experiment 1 examined the characteristics that are associated with the appeal of icons. Given the remarkable correspondences between what is known about the stimulus characteristics that correlate with both performance and appeal, it was hypothesized that icon characteristics which affected user performance would have similar effects on appeal, i.e. that correlations between icon characteristics and appeal would be similar to those observed between icon characteristics and performance. This was indeed the case. First, icon familiarity accounted for the largest proportion of the variance in appeal. Second, icon concreteness was a less effective predictor of appeal than familiarity. Third, icon concreteness did not predict appeal once the effects of icon familiarity were accounted for. Fourth, visual complexity accounted for a significant amount of the variance although this overlapped to some degree with icon familiarity.

3 Experiment 2: Processing fluency effects in an icon search task

If the processing fluency hypothesis is correct, then icons that are likely to be easier to process (i.e., familiar and simple icons) should attract more positive evaluations. In Experiment 2 participants carried out a search task designed to be analogous to everyday search for icons on an interface. The effect of icon familiarity and visual complexity on search performance was examined by varying the familiarity and complexity of the icons presented in the task orthogonally to create 4 types of icon; familiar and simple, familiar and complex, unfamiliar and simple, unfamiliar and complex (see Figure 4). The appeal of all 4 types of icons was held constant (see Materials and Table 2 for details). This meant that any differences in appeal emerging during the study could be attributed to differences in performance arising during the search task, rather than merely being the result of prior appeal differences between experimental materials. In order to ensure that appeal judgements were not simply the result of initial impressions, icons were presented to participants over a series of 9 blocks of experimental trials and ratings of liking were obtained after the first and the ninth blocks of trials. In this way Experiment 2 was designed to provide a strong test of the processing fluency hypothesis, examining the possibility that icon appeal “is a function of the perceiver’s processing dynamic: The more fluently the perceiver can process an object, the more positive [their] aesthetic response.” (Reber et al., 2004, p. 366).

On the basis of previous research (McDougall et al., 2000; Scott, 2003) we expected that visual search for simple icons would be easier than for complex icons (i.e. a main effect

of complexity). We also expected that familiar icons would be located in the array more easily than unfamiliar icons since search is partially driven by our prior knowledge of stimuli (i.e. a main effect of familiarity; Liu, 2005; Wang et al., 1994, Wolfe & Alvarez, 2011). Finally, as participants gained experience with the search task and the icon set, we expected that processing would become easier (i.e., a main effect of experience). It was expected this ease of processing would be reflected in measures of performance (search times) and appeal evaluations (subjective ratings of liking).

3.1 Method

3.1.1 Participants

Forty-two undergraduate and postgraduate students from Bournemouth University took part in this experiment. Twenty-one participants provided ratings of liking after completing 1 block of experimental trials. Their mean age was 20.83 years (12 females, 9 males). Twenty-one participants provided ratings of liking after completing 9 blocks of experimental trials. Their mean age was 22.33 years (13 females, 8 males). Each participant received £12 or course credits for taking part in the study.

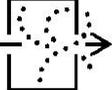
Familiarity	Visual Complexity	
	Simple	Complex
Familiar		
	Picnic Area	Risk of explosion
Unfamiliar		
	Balance	Rinse

Figure 4: Types of icons presented in the search and identification tasks in Experiments 2-4. The visual complexity and familiarity of icons was varied orthogonally while appeal was held constant.

3.1.2 Materials

Icons were selected from a corpus of 239 icons rated for familiarity, complexity and appeal (McDougall et al., 1999 and Experiment 1). As shown in Figure 4, the 40 icons presented in the search trials varied orthogonally in their rated familiarity and visual complexity. There were (a) 10 familiar and complex icons, (b) 10 familiar and simple icons, (c) 10 unfamiliar and complex, and (d) 10 unfamiliar and simple icons. Using ratings from the corpus, icons were selected which had previously been rated independently as either relatively visually complex or simple and either familiar or unfamiliar. Based on appeal ratings obtained in Experiment 1, the icons selected *did not differ* in their perceived appeal. A series of one-way analyses of variance followed by Newman-Keuls comparisons was

Table 2: Mean ratings (and standard deviations) of visual complexity, familiarity, and appeal for each icon type and the results of one-way analyses and Newman-Keuls comparisons examining differences between stimulus types used in Experiments 2 and 3. All ratings were on a 1-5 scale, with 5 representing a high value of the characteristic concerned.

Icon characteristics	Complex Familiar (CF)	Complex Unfamiliar (CUF)	Simple Familiar (SF)	Simple Unfamiliar (SUF)	F-value	Newman-Keuls comparisons
Visual complexity	3.44 (0.38)	3.47 (0.34)	2.00 (0.35)	2.02 (0.21)	F(3,36)=63.88, p<.001	CF=CUF > SF=SUF
Familiarity	3.47 (0.38)	2.02 (0.34)	3.44 (0.35)	2.00 (0.21)	F(3,36)=131.68, p<.001	CF=SF > CUF=SUF
Appeal	2.96 (0.37)	2.68 (0.48)	3.29(0.43)	2.90 (0.35)	F(3,36)=2.69, p>.05	CF=CUF=SF=SUF

conducted to ensure that icons differed in accordance with the requirements of each experimental condition (see Table 2):-

- (i) Ratings of complexity differed significantly; simple familiar and unfamiliar icons had significantly lower complexity ratings than complex familiar and unfamiliar icons
- (ii) Familiar simple and complex icons had higher familiarity ratings than unfamiliar simple and complex icons.
- (iii) Ratings of icon appeal did *not* differ significantly across experimental conditions.

3.1.3 Procedure

3.1.3.1 *Search task.* Participants searched for a pre-determined target icon in an array of nine icons (see Figure 5). The procedure for each experimental trial was as follows:-

- (a) Participants were initially shown the target for 2 seconds.
- (b) An OK button appeared following target offset. Participants then pressed the OK button to display the array. This ensured participants started each trial with the mouse pointer at the same point on the display.
- (c) An array of 9 icons appeared after the OK button was depressed. Participants used the mouse to click as quickly as possible on the target icon. Once participants had clicked on an icon in the array, the array disappeared.
- (d) There was an inter-stimulus interval of 1s after icon array offset and then the next experimental trial began.

There were 9 blocks of trials, with 40 trials per block. In each block, 40 icons were targets, 10 icons of each icon type (shown in Figure 4). Icons in the set appeared equally often as targets and distractors. Performance was measured by the time participants took to search for an icon in an array (a task similar to searching for icons on an interface) and by search accuracy. Accuracy scores were therefore the number of each type of icon correctly matched with an icon in the array during search (i.e. a score out of 10 for each icon type).

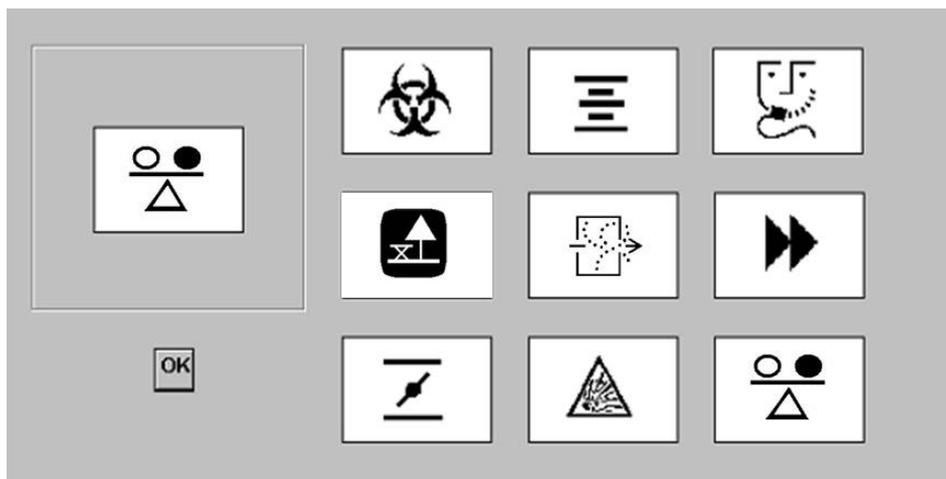


Figure 5: The search task used in Experiment 2 (see *Procedure* for details)

3.1.3.2 *Appeal rating task.* Participants were asked to rate icons on a 1-5 scale (from really dislike to really like) *either* after completing 1 block of experimental trials *or* after completing 9 blocks of experimental trials. Icons were randomized prior to rating and presented in booklets to participants.

3.1.4 Design

Participants completed two tasks, a search task and a liking rating task. The search task employed a 2 (Visual Complexity: simple vs. complex) x 2 (Familiarity: familiar vs. unfamiliar) x 9 (Block: 1-9) mixed design. The Complexity by Familiarity manipulation yielded 4 icon types and there were 10 icons of each type. There were 40 trials in each the 9 blocks of trials. The dependent measure for the search task was the time taken to find the target icon in an array of icons.

The rating task employed a 2 (Visual Complexity: simple vs. complex) x 2 (Familiarity: familiar vs. unfamiliar) x 2 (Experience: block 1 vs. block 9) design where Visual Complexity and Familiarity were repeated measures but Experience was manipulated between-participants. The dependent measure for this task was participants' subjective ratings of liking using a 1-5 scale.

3.2 Results and Discussion

3.2.1 Accuracy

Participants' accuracy was very high throughout, $M=94.33\%$, $SD=4.86\%$. For this reason accuracy data has not been analysed further.

3.2.2 Response Times

A repeated-measures analysis of variance (ANOVA) examined the effects of experience (block 1 vs. block 9), icon familiarity (familiar vs. unfamiliar) and icon complexity (simple vs. complex) on participants' response times. A summary of main effects and interactions on response times is shown in Table 3. Response times reduced as participants gained experience with icons in the search task, $F(8,328) = 13.17$, $p < .001$, $\eta^2 = .243$, $M(\text{block 1})=1239$ ms, $SD = 194$ ms; $M(\text{block9})=1101$ ms, $SD = 133$ ms. Response times were faster for familiar in comparison to unfamiliar icons, $F(1,41) = 80.76$, $p < .001$, $\eta^2 = .663$; $M(\text{familiar}) = 1112$ ms, $SD = 129$ ms; $M(\text{unfamiliar}) = 1174$ ms, $SD = 136$ ms. Response times were faster for simple in comparison to complex icons, $F(1,41) = 90.34$, $p < .001$, $\eta^2 = .696$; $M(\text{simple}) = 1105$, $SD = 135$ ms; $M(\text{complex}) = 1181$ ms, $SD = 132$ ms. There was a significant interaction between familiarity and complexity, $F(1,41) = 10.69$, $p = .002$, $\eta^2 = .207$. This interaction is shown in Figure 6a. Further analyses revealed that the difference in response times between familiar and unfamiliar icons was greater for simple icons than for complex icons, $t(41) = 3.27$, $p = .002$. No other interactions were significant.

3.2.3 Ratings of appeal

A repeated-measures ANOVA examined the effects of experience (block 1 vs. block 9), icon familiarity (familiar vs. unfamiliar) and icon visual complexity (simple vs. complex) on participants' ratings of liking. A summary of main effects and interactions on liking ratings is shown in Table 3. There was a significant main effect of experience, $F(1,40) = 5.46$, $p = .025$, $\eta^2 = .120$. Participants who had gained more experience with the icons gave them higher ratings of liking, $M(\text{block 9})=3.06$, than those with less experience of the icons, $M(\text{block 1})=2.81$. Ratings were higher for familiar than for unfamiliar icons, $F(1,40) = 17.55$, $p < .001$, $\eta^2 = .305$ and for simple as opposed to complex icons, $F(1,40) = 8.87$, $p < .001$, $\eta^2 = .219$.

= .180. Furthermore, there was a significant interaction between familiarity and complexity, $F(1,40) = 4.77, p = .035, \eta^2 = .107$ (see Figure 6b). A paired t-test was carried out to examine the interaction further and showed that the difference in ratings between familiar and unfamiliar icons was greater for simple icons than for complex icons, $t(41) = 2.17, p = .035$. No other interactions were significant.

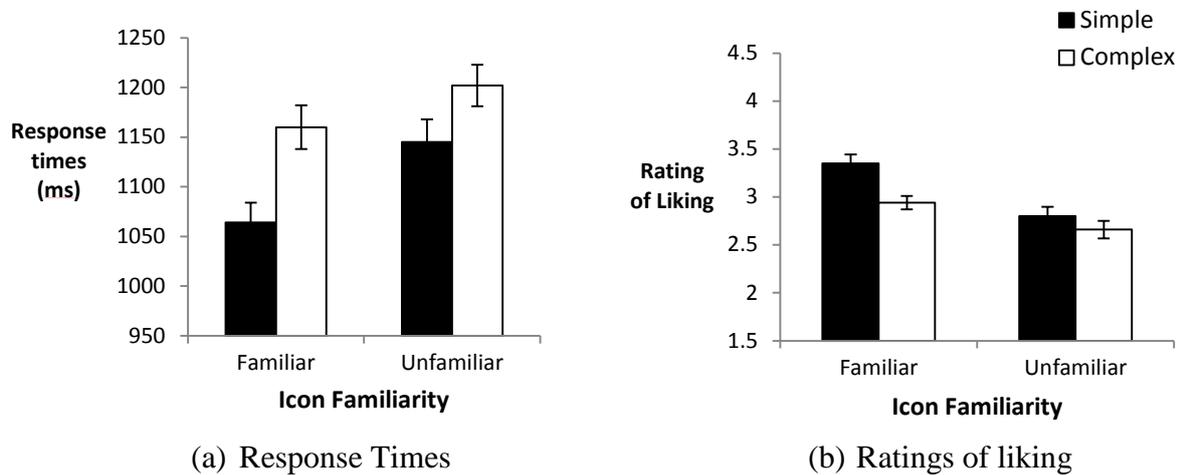


Figure 6: Experiment 2. Interactions between icon familiarity and complexity. Means and standard errors for (a) response times and (b) ratings of liking in the search task

In this experiment icons which were processed more efficiently were expected to produce an implicit experience of fluency which would be attributed to greater liking leading to more positive evaluations of appeal for these icons. The findings for ratings of liking reflect those found for response times; simple and familiar icons attracted higher ratings of liking. When combined, these properties appeared to have particular benefits, with familiar simple icons resulting in the lowest response times and attracting the highest ratings of liking. Ratings of liking also differed based on experience; when ratings of liking were obtained earlier in learning trials at block 1 ratings were lower than at block 9 reflecting the reduction in response time over blocks of trials.

The results of Experiment 2 showed that ratings of liking mirrored the pattern of response times (see Figure 6 and Table 3). Although, the significant main effects of familiarity or visual complexity of icons in the appeal ratings *could* be attributed to the effects those characteristics are already known to have on appeal judgements (e.g. Experiment 1; Bauerly & Liu, 2008; Moshagen & Theilich, 2010; Monahan et al., 2001; Zajonc, 2001), the interaction between icon familiarity and visual complexity can only be attributed to the relative ease or difficulty of processing emerging *during* the visual search task. These results suggest a dynamic relationship between performance and appeal that is mediated by processing fluency.

Table 3: Summary of experimental findings. Shaded cells indicate *predicted* main effects and interactions in performance measures (RT and accuracy) and appeal ratings. Predicted effects and interactions were different between Experiment 2 (search task) and Experiments 3 & 4 (search-and-identification task). Ticks and crosses denote the effects *observed*: ticks (✓) denote significant effects; crosses (✗) denote non-significant effects; dashes (-) indicate insufficient data to conduct analyses.

Experimental effects	RT	Accuracy	Appeal
Experiment 2			
<i>Icon characteristics</i>			
Visual complexity: simple vs complex icons	✓	-	✓
Familiarity: familiar vs unfamiliar icons	✓	-	✓
Familiarity x Visual complexity	✓	-	✓
<i>Experience (learning during blocks of experimental trials)</i>			
Experience (block 1 vs block 9)	✓	-	✓
Experience x Visual complexity	✗	-	✗
Experience x Familiarity	✗	-	✗
Experience x Visual complexity x Familiarity	✗	-	✗
Experiment 3			
<i>Icon characteristics</i>			
Visual complexity: simple vs complex icons	✓	✗	✓
Familiarity: familiar vs unfamiliar icons	✓	✓	✓
Familiarity x Visual complexity	✓	✓	✓
<i>Experience (learning during blocks of experimental trials)</i>			
Experience (block 1 vs block 9)	✓	✓	✗
Experience x Visual complexity	✗	✗	✗
Experience x Familiarity	✓	✓	✗
Experience x Visual complexity x Familiarity	✓	✗	✗
Experiment 4			
<i>Icon characteristics</i>			
Visual complexity: simple vs complex icons	✓	✗	✓
Familiarity: familiar vs unfamiliar icons	✓	✓	✓
Familiarity x Visual complexity	✓	✓	✓
<i>Experience (learning during blocks of experimental trials)</i>			
Experience (block 1 vs block 9)	✓	✓	✗
Experience x Visual complexity	✗	✗	✗
Experience x Familiarity	✓	✓	✓
Experience x Visual complexity x Familiarity	✓	✗	✓

4 Experiment 3: Processing fluency effects in an icon identification task

In Experiment 2 the pattern of liking ratings closely mirrored the pattern of performance, and specifically response times. It was suggested that the similarity between the pattern of liking and the pattern of performance may be mediated by processing fluency rather than icon characteristics alone. An even stronger test of the processing fluency hypothesis as a dynamic mediator of liking ratings would be to change the task (so changing ease of processing) while keeping the icons used the same. This is exactly what was done in Experiment 3. Participants were asked to carry out an icon identification task but the icon set was the same as used in the search task in Experiment 2. Participants were shown the function of the icon and then asked to identify the matching icon. This meant that participants had to search for a semantic rather than a visual match in the array. Critically, patterns of performance in the identification task are different from the search task employed in Experiment 2 (McDougall et al., 2000). Specifically, in this kind of task the expected pattern of results is as follows (see also Table 3):-

- (i) When icons are unfamiliar, the icon-function associations need to be learned, so in early learning trials unfamiliar icons will be identified less accurately than in later trials. When unfamiliar icon-function associations have been learned (by Block 9) there should be no differences in accuracy between icon types (i.e. for accuracy scores there should be an interaction between icon familiarity and experience over blocks of trials).
- (ii) Familiar icons will also be identified faster in comparison to unfamiliar icons. This will be particularly apparent in earlier trials although the difference in response times between familiar and unfamiliar icons should diminish as icon-function associations are learned in later trials (i.e. there should be an interaction between icon familiarity and experience in response times).
- (iii) Visual complexity may also have an effect on identification *times* since prior knowledge of shapes and objects is known to affect search times (i.e. a main effect of visual complexity on response times).
- (iv) Visual complexity has no effect on identification *accuracy* because it is likely to be less important compared to our familiarity with icons and the icon-function associations.

If appeal evaluations are determined by ease of processing in the identification task, despite the icons used in Experiment 3 being identical to those used in Experiment 2, this would provide the strongest evidence so far for processing fluency as a dynamic mediator of appeal. It would show not only evaluations of icon appeal mirror performance (as we showed in Experiment 2), but also that the pattern of liking ratings for the *same* icons *changes* when performance changes as a result of using a different task which yields a different pattern of performance (see Table 3 for the predicted pattern of findings, which are shaded in grey). If, on the other hand, appeal is determined primarily by the nature of the stimuli, rather than processing fluency per se, then we would not expect to see any change in appeal evaluations from those observed in Experiment 2.

4.1 Method

4.1.1 Participants

Thirty-eight undergraduate and postgraduate students from Bournemouth University took part in this experiment. Twenty-one participants provided ratings of liking after completing 1 block of experimental trials. Their mean age was 21.44 years, $SD=1.39$ (11 females, 10 males). Seventeen participants provided ratings of liking after completing 9 blocks of experimental trials. Their mean age was 22.85 years, $SD=6.65$ (9 females, 7 males). Each participant received course credits or £8 for taking part in the study.

4.1.2 Materials

The materials were identical to those employed in Experiment 2. Familiarity and visual complexity were varied orthogonally while the appeal of the icons prior to the study was held constant (see Table 2).

4.1.3 Procedure

4.1.3.1 Identification task. In Experiment 3, participants were shown the written function of the icons and asked to identify the matching icon in the 9-icon array. Corrective feedback was given following errors, to facilitate learning of the correct icon-function associations. The procedure for each experimental trial was as follows:

- (a) Participants were shown the icon function for 2 seconds (e.g., the word “biohazard”).
- (b) An OK button appeared after the function offset. Participants then pressed the OK button to display the array.
- (c) Participants used the mouse to select the icon which they thought matched the function. If the correct icon was chosen, the array disappeared and the next experimental trial began. When an error was made, participants were shown the array twice more, if required, so that they could make another attempt to match the function to its icon. If, on their third attempt, participants were still unable to choose the correct icon match, the array disappeared leaving the correct icon. This provided participants with feedback so that they could gradually learn the icon-function associations.

As in Experiment 2 there were 9 blocks of trials with 40 trials per block. The function name for each of the 40 icons was presented once in each block of trials. Accuracy was measured as the number of icons correctly identified at the first attempt in each trial. For each block of 40 trials, a maximum accuracy score of 10 could be obtained for each of the 4 types of icons presented.

4.1.3.2 Liking rating task. Participants were asked to rate icons on a 1-5 scale after completing either 1 or 9 blocks of experimental trials in exactly the same manner as Experiment 2.

4.1.4 Design

The experimental design was identical to Experiment 2, except that participants were asked to identify icons in an array given the function of the icon, rather than being shown the target icon itself.

4.2 Results and Discussion

4.2.1 Accuracy and Response Times

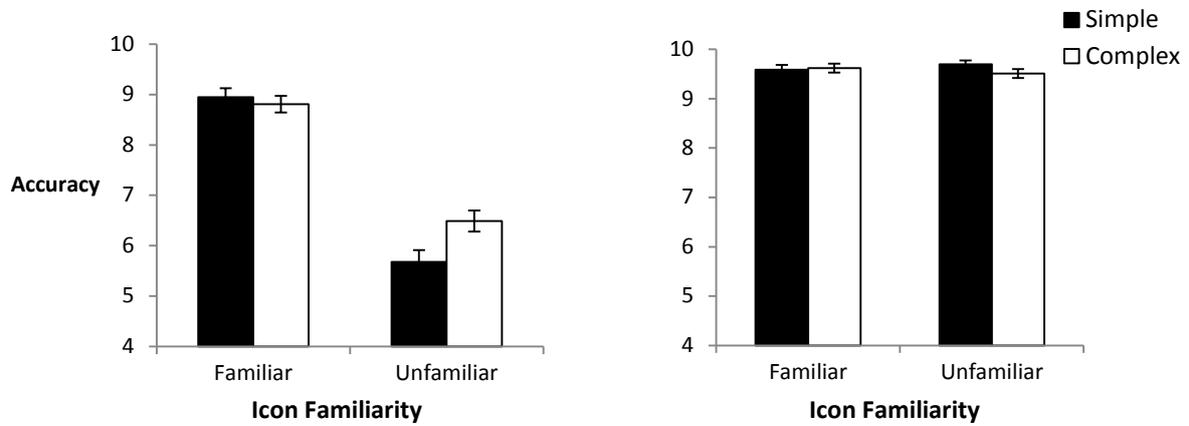
Two repeated-measures ANOVAs examined the effects of (blocks of trials 1-9), icon familiarity (familiar vs. unfamiliar) and icon visual complexity (simple vs. complex) on participants' accuracy and response times. A summary of main effects and interactions on icon identification accuracy, response times, and liking ratings for Experiment 3 is shown in Table 3 and Figure 7.

4.2.2.1 Accuracy. Accuracy increased as participants gained experience with icons in the identification task, $F(8,288) = 101.93$, $p < .001$, partial $\eta^2 = .739$, M (Block1) = 7.48, $SD = 0.70$; M (Block 9) = 9.63, $SD = 0.26$. Responses were more accurate for familiar icons, $M=9.54$, $SD = 0.30$, than for unfamiliar icons, $M=8.92$, $SD = 0.47$, $F(1,36) = 74.18$, $p < .001$, partial $\eta^2 = .673$. The effects of familiarity reduced over time and this was apparent in a significant interaction between experience and familiarity, $F(8, 288) = 49.68$, $p < .001$, partial $\eta^2 = .580$. Paired-samples t-tests were carried out to examine differences between familiar and unfamiliar icons in Block 1 and Block 9. The criterion p-value was adjusted to .025 using the Bonferroni correction to adjust for family-wise error rate. These comparisons revealed a significant difference in accuracy between familiar and unfamiliar icons at Block 1, $t(36) = 13.21$, $p < .001$ but not at Block 9, $t(36) = -.61$, $p = .544$.

Visual complexity, in contrast, did not affect accuracy, $F(1,36) = 1.39$, $p = .246$, partial $\eta^2 = .037$, M (simple) = 9.20, $SD = 0.34$; M (complex) = 9.25, $SD = 0.36$, and, overall, its effects did not change as a result of experience, $F(8, 288) = 1.79$, $p = .078$, partial $\eta^2 = .047$. As in Experiment 2, there was a significant interaction between the effects of familiarity and complexity, $F(1,36) = 6.96$, $p = .012$, partial $\eta^2 = .162$. Although the interaction between these effects appeared to reduce somewhat as participants gained experience with the task the 3-way interaction between experience, icon familiarity and icon complexity was not significant, $F(8,288) = 1.97$, $p = .05$, partial $\eta^2 = .052$. However, for the purpose of comparison with Experiment 2, the effects of familiarity and complexity on response times initially (at Block 1) and at the end of the identification task (at Block 9) are illustrated in Figure 5(a) and (b). T-tests were carried out to examine differences between simple and complex familiar and unfamiliar icons in Block 1 and Block 9. The criterion p-value was adjusted to .0125 using the Bonferroni correction to adjust for family-wise error rate. These comparisons revealed a significant difference in accuracy for simple and complex unfamiliar icons at Block 1, $t(36) = 2.61$, $p = .006$, but no other differences between simple and complex icons were apparent.

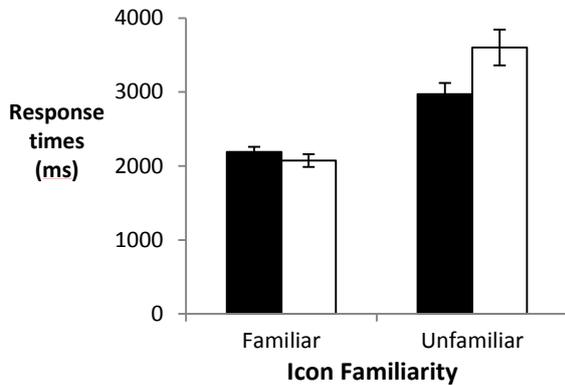
4.2.2.2 Response Times. Response times reduced over blocks of trials as participants gained experience with the identification task, $F(8,288) = 100.87$, $p < .001$, partial $\eta^2 = .737$, M (Block 1) = 2708ms, $SD=677ms$, M (Block 9) = 1403ms, $SD = 292ms$. Familiarity also affected response times, $F(1,36) = 141.40$, $p < .001$, partial $\eta^2 = .797$; M (familiar) = 1429ms, $SD = 291ms$; M (unfamiliar) = 1894ms, $SD = 381ms$. The effects of familiarity reduced with experience, $F(8, 288) = 24.44$, $p < .001$, partial $\eta^2 = .404$. T-test comparisons using the Bonferroni correction revealed that while response time differences between familiar and unfamiliar icons reduced with experience, they remained significant throughout (Block 1: $t(36) = -7.68$, $p < .001$; Block 9: $t(36) = -5.45$, $p < .001$).

Figure 7: Experiment 3. Interactions between icon familiarity and complexity. Means and standard errors for accuracy after (a) 1 block and (b) 9 blocks of trials; response times after (c) 1 block and (d) 9 blocks of trials; and ratings of liking after (e) 1 block and (f) 9 blocks of trials.

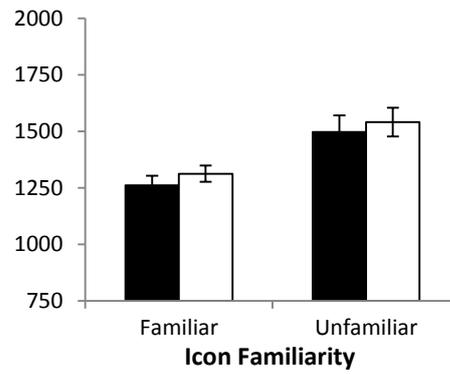


(a) Accuracy Block 1

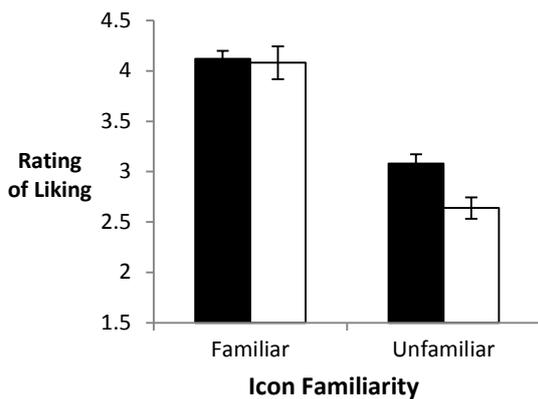
(b) Accuracy Block 9



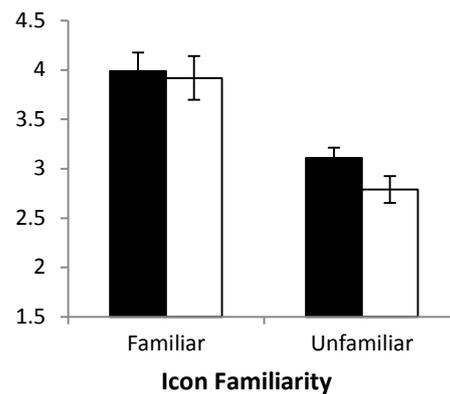
(c) Response Times Block 1



(d) Response Times Block 9



(e) Liking Ratings Block 1



(f) Liking Ratings Block 9

Complex icons took longer to identify than simple icons, $M(\text{simple}) = 1628\text{ms}$, $SD = 307\text{ms}$, $M(\text{complex}) = 1694\text{ms}$, $SD = 293\text{ms}$, $F(1,36) = 5.54$, $p = .04$, partial $\eta^2 = .133$, however, this difference did not reduce as a result of experience, $F(8, 288) = 1.62$, $p = .118$, partial $\eta^2 = .043$. There was a significant interaction between the effects of familiarity and complexity, $F(1,36) = 10.01$, $p = .003$, partial $\eta^2 = .218$ (see Figure 5). The interaction between these effects reduced as participants gained experience across blocks of trials and resulted in a significant 3-way interaction, $F(8,288) = 4.82$, $p < .001$, partial $\eta^2 = .118$. Two 2-way repeated-measures ANOVAs were carried out to examine the 3-way interaction. For Block 1, the visual complexity x familiarity interaction was also significant, $F(1,36) = 15.94$, $p < .001$; partial $\eta^2 = .307$. Paired-samples t-tests with the p-value adjusted to .0125 to correct for the family-wise error rate using the Bonferroni correction showed significantly lower response times for simple unfamiliar icons compared to complex unfamiliar icons, $t(36) = -3.28$, $p = .002$, but no difference between simple and complex familiar icons, $t(36) = 1.42$, $p = .163$. For Block 9, there was no significant interaction, $F(1,36) = 1.11$, $p = .299$. The effects of familiarity and complexity initially (at Block 1) and at the end of the identification task (at Block 9) are illustrated in Figure 7(c) and (d).

4.2.2.3 Ratings of liking. A mixed 3-way ANOVA was used to examine differences in appeal ratings as a result of experience (block 1 vs. block 9) as the between-participants factor and icon familiarity (familiar vs. unfamiliar) and icon visual complexity (simple vs. complex) as the within-participants factors.

Appeal ratings did not differ as a result of experience, $F(1,36) = 0.710$, $p = .409$, partial $\eta^2 = .034$, $M(\text{block 1}) = 3.48$, $SD = 0.29$, $M(\text{block 9}) = 3.46$, $SD = 0.48$. However, appeal ratings differed as a result of icon familiarity, $F(1,35) = 117.73$, $p < .001$, partial $\eta^2 = .771$, $M(\text{familiar}) = 4.03$, $SD = 0.58$; $M(\text{unfamiliar}) = 2.91$, $SD = 0.40$, and icon complexity, $F(1,35) = 16.65$, $p < .001$, partial $\eta^2 = .332$, $M(\text{simple}) = 3.58$, $SD = 0.40$; $M(\text{complex}) = 3.36$, $SD = 0.44$. There was also an interaction between familiarity and complexity, $F(1,35) = 9.73$, $p < .001$, partial $\eta^2 = .217$. Paired-samples t-tests revealed that ratings for unfamiliar simple and complex icons differed significantly, $t(36) = -4.85$, $p = .000$, while familiar simple and complex icon ratings did not, $t(36) = .90$, $p = .371$. The effects of familiarity and complexity initially (at Block 1) and at the end of the identification task (at Block 9) are illustrated in Figure 5(e) and (f). Bonferroni corrected paired-samples t-tests revealed significant differences in ratings of liking between unfamiliar complex and simple icons at Block 1, $t(19) = -4.15$, $p = .001$, but not in Block 9, $t(16) = -2.19$, $p = .022$. There were no significant differences between familiar complex and simple icons at either Block 1 or 9. None of the interactions with experience were significant.

Findings from Experiment 3 are illustrated in Figure 7 (see also Table 3). They can be summarized as follows:-

(i) *Accuracy*

Identification accuracy was much poorer for unfamiliar icons in Block 1. However, by Block 9 icon-function associations had been learned, reducing any differences in accuracy between icon types. As expected, visual complexity had little overall effect on identification accuracy but, as Figure 7 shows, the extra detail provided in complex icons may have helped participants make inferences initially about icon-function relationships.

(ii) Response Times

Visual search times, however, were generally faster for simple icons. In addition, familiar icons were identified more quickly than unfamiliar icons; this was particularly apparent in earlier trials, though differences remained even in later trials.

(iii) Evaluations of Appeal

Ratings of liking appeared to mirror response times more than accuracy (see Figure 7). Liking ratings differed significantly as a result of both icon complexity and familiarity. Differences in ratings between familiar icons and unfamiliar icons remained, even in later trials at Block 9. In this instance attributions therefore appeared to mirror speed, rather than accuracy, of responses.

So, does changing the task – and the nature of the cognitive processing required – change ratings of liking as indicated in Figure 1? By and large, this appears to have been the case with respect to search times. However, if appeal ratings were to fully reflect changes in performance, then the effects of familiarity would reduce as participants learned the icon-function relationships across blocks of trials (i.e. there would be an Experience x Familiarity interaction, see areas shaded in grey in Table 3). This was not the case for appeal evaluations. One possible reason for this might be that ratings of liking at Blocks 1 and 9 were obtained from different groups of participants, reducing the likelihood of using learning of the icon-function relationship as a heuristic cue. In order to explore this possibility, a further experiment was conducted replicating Experiment 3 but ratings of liking were obtained *twice* from each participant, once after completion of 1 block and then again after 9 blocks of trials. The aim of this manipulation was to give greater emphasis, albeit implicitly, to participants' growing experience with the icons and the task over experimental trials.

5 Experiment 4: Does learning icon-function relationships act as a cue in evaluations of appeal?

If ease, or dis-ease, of processing acts as a heuristic cue in determining appeal evaluations then it should be possible to map the pattern of performance in a given task to the pattern of appeal evaluations. In Experiment 3 this was largely, but not wholly, the case. In contrast to accuracy and response times, differences between appeal evaluations for familiar and unfamiliar icons did *not* diminish as icon-function relationships were learned over 9 blocks of trials. One possible explanation is that, because ratings of liking were obtained from different groups of participants after completion of *either 1 or 9* blocks of trials participants were more aware of stimulus cues such as icon familiarity and complexity when rating liking rather than relative changes in ease of processing resulting from greater knowledge of the task and icon-function associations. Thus, when asked to rate liking, participants were less aware of, or discounted, their learning between blocks of trials given the greater salience of icon familiarity and complexity. *Discounting* of possible heuristic cues appears to occur in situations where multiple cues are available because individuals attribute fluency to those which are most obvious and available (see Alter & Oppenheimer, 2009; Alter, Oppenheimer, Epley & Eyre, 2007; Belke, Leder, & Carbon, 2015; Oppenheimer, 2004, for discussions of heuristic discounting). Experiment 4 explored this possibility by replicating Experiment 3 but with ratings of liking obtained twice from the same participants, once after completion of

Block 1 and a second time after completion of Block 9. The aim of this manipulation was to implicitly highlight *gaining experience with the icons* over blocks of trials.

5.1 Method

5.1.1 Participants

Twenty-one undergraduate and postgraduate students from Bournemouth University took part in this experiment. Their mean age was 20 years 4 months, $SD=5$ months (20 females, 1 male). Each participant received course credits for taking part in the study.

5.1.2 Procedure & Materials

The procedure and materials were identical to Experiment 3, with the exception that in Experiment 4 participants rated the appeal of the icons which had appeared in the array after completing 1 *and* after completing 9 blocks of experimental trials.

5.1.3 Design

The design for the identification task was identical to Experiment 3. The appeal evaluation task employed a 2 (Familiarity: familiar vs. unfamiliar) \times 2 (Visual complexity: simple vs. complex), \times 2 (Experience: Block 1 vs. 9) repeated-measures design. Thus, unlike Experiment 3, in Experiment 4 Experience was a within-participant factor. Accuracy and response time were the dependent measures for the identification task, and appeal ratings were the dependent measure for the appeal evaluation task.

5.2 Results and Discussion

The pattern of findings in this experiment was very similar to Experiment 3 (c.f. Figures 7 and 8).

5.2.1 Accuracy and Response times

Repeated measures ANOVAs were conducted to examine the effects of experience (blocks of trials 1-9), icon familiarity (familiar vs. unfamiliar) and icon visual complexity (simple vs. complex) on participants' accuracy, response times and liking ratings. A summary of main effects and interactions for each dependent measure appears in Table 3.

5.2.1.1 Accuracy. Accuracy increased across blocks of trials as participants gained experience with the identification task, $F(8,160) = 114.76$, $p < .001$, partial $\eta^2=.852$; M (Block1) = 7.05, $SD = 0.66$; M (Block 9) = 9.62, $SD = 0.36$. Familiar icons were identified more accurately than unfamiliar icons, $F(1,20) = 71.50$, $p < .001$, partial $\eta^2=.781$; M (familiar) = 9.53, $SD = 0.30$; M (unfamiliar) = 8.91, $SD = 0.17$. The effects of familiarity reduced across blocks of trials as users learned icon-function associations, $F(8,160) = 29.23$, $p < .001$, partial $\eta^2=.594$. As can be seen from Figure 8, there was a large difference in accuracy scores between familiar and unfamiliar icons in the first block of trials, $t(21)=8.77$, $p<.001$, but not at block 9, $t < 1$. Visual complexity did not affect accuracy, $F(1,20) = 0.24$, $p = .625$, partial $\eta^2=.012$; M (simple) = 9.20, $SD = 0.22$; M (complex) = 9.24, $SD = 0.27$. No other interactions with complexity were significant. For the purpose of comparison with Experiments 2 and 3, the effects of familiarity and complexity on accuracy initially at Block 1 and later at Block 9 are illustrated in Figure 8(a) and (b).

5.2.1.2 Response Times. Response times reduced over blocks of trials as participants gained experience with the task, $F(8,160) = 58.52, p < .001$, partial $\eta^2 = .745$; $M(\text{Block 1}) = 2512\text{ms}$, $SD = 592\text{ms}$; $M(\text{Block 9}) = 1368\text{ms}$, $SD = 155\text{ms}$. Familiar icons were identified more quickly, $F(1,20) = 84.26, p < .001$, partial $\eta^2 = .808$; $M(\text{familiar}) = 1397\text{ms}$, $SD = 179\text{ms}$; $M(\text{unfamiliar}) = 1742\text{ms}$, $SD = 265\text{ms}$. The effects of familiarity reduced over time as participants learned icon-function associations, $F(8,160) = 16.02, p < .001$, partial $\eta^2 = .455$. T-tests using the Bonferroni correction revealed that while response time differences between familiar and unfamiliar icons reduced with experience, they remained significant throughout (Block 1: $t(21) = -6.16, p < .001$; Block 9, $t(21) = 3.74, p < .001$).

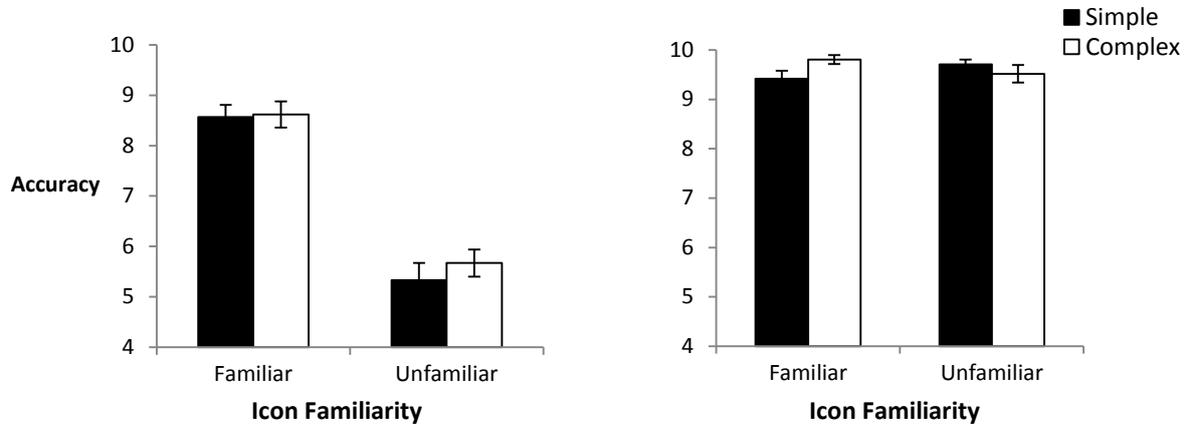
Complex icons took longer to identify than simple icons, $F(1,20) = 8.09, p = .010$, partial $\eta^2 = .288$, $M(\text{simple}) = 1540\text{ms}$, $SD = 219\text{ms}$; $M(\text{complex}) = 1599\text{ms}$, $SD = 209\text{ms}$. Familiarity and complexity interacted significantly, $F(1,20) = 10.76, p = .004$, partial $\eta^2 = .350$ (see Figure 6c and d). There was a significant 3-way interaction between experience, icon familiarity and icon complexity, $F(8,160) = 4.62, p < .001$, partial $\eta^2 = .188$. Two 2-way repeated-measures ANOVAs were carried out to examine the 3-way interaction. At Block 1, the visual complexity x familiarity interaction was significant, $F(1,20) = 10.20, p = .005$; partial $\eta^2 = .338$. At Block 9, the interaction was not significant, $F(1,20) = 1.11, p = .299$. Paired-samples t-tests were carried out to examine these effects further with the criterion p-value adjusted to .0125 using the Bonferroni correction to correct for the family-wise error rate. A significant difference in response time was observed between simple and complex icons for unfamiliar icons at Block 9, $t(20) = 3.95, p = .001$, but not for other pairwise comparisons of simple and complex icons. The effects of familiarity and complexity initially (at Block 1) and at the end of the identification task (at Block 9) are illustrated in Figure 6(c) and (d).

5.3.1 Ratings of liking

Appeal ratings did not differ between blocks 1 and 9, $F(1,20) = 0.71, p = .409$, partial $\eta^2 = .034$, $M(\text{Block 1}) = 3.23, SD = 0.65$; $M(\text{Block 9}) = 3.42, SD = 0.79$. Familiar icons were rated more highly than unfamiliar icons, $F(1,20) = 150.97, p < .001$, partial $\eta^2 = .883$, $M(\text{familiar}) = 3.98, SD = 0.64$; $M(\text{unfamiliar}) = 2.76, SD = 0.65$. The effects of familiarity on ratings of liking reduced between Block 1 and Block 9 producing a significant interaction between familiarity and experience, $F(1,20) = 17.99, p < .001$, partial $\eta^2 = .474$.

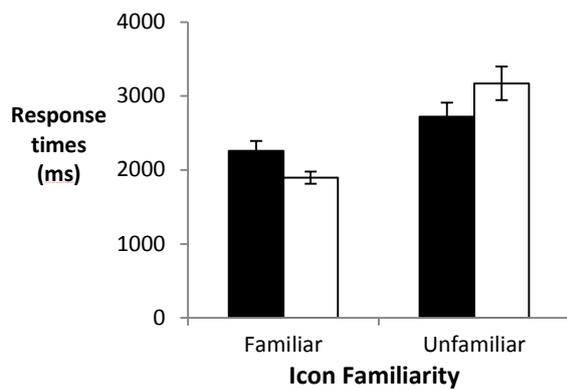
Simple icons were rated more highly than complex icons $F(1,20) = 34.96, p < .001$, partial $\eta^2 = .636$, $M(\text{simple}) = 3.56, SD = 0.64$; $M(\text{complex}) = 3.18, SD = 0.60$, and there was a significant 3-way interaction between experience, icon familiarity and icon complexity, $F(1,20) = 34.94, p < .001$, partial $\eta^2 = .636$. Two 2-way repeated-measures ANOVAs were carried out to examine the 3-way interaction. The visual complexity x familiarity interaction was significant at both Block 1, $F(1,20) = 32.80, p < .001$; partial $\eta^2 = .621$, and Block 9, $F(1,20) = 12.45, p = .002$, partial $\eta^2 = .384$. Paired-samples t-tests were carried out to examine these effects further using the Bonferroni correction with the criterion p-value adjusted to .0125. Ratings at Block 1 differed between simple and complex icons for unfamiliar icons, $t(20) = 7.95, p < .001$, but not for familiar icons, $t(20) = .192, p = .850$.

Figure 8: Experiment 4. Interactions between icon familiarity and complexity. Means and standard errors for accuracy after (a) 1 block and (b) 9 blocks of trials; response times after (c) 1 block and (d) 9 blocks of trials; and ratings of liking after (e) 1 block and (f) 9 blocks of trials.

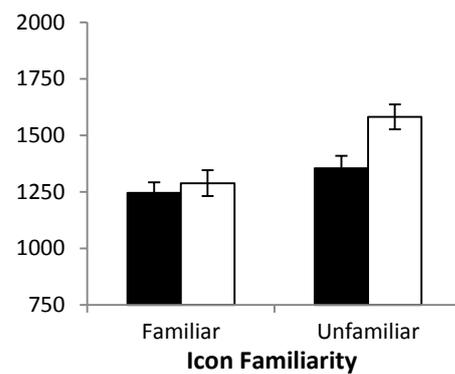


(a) Accuracy Block 1

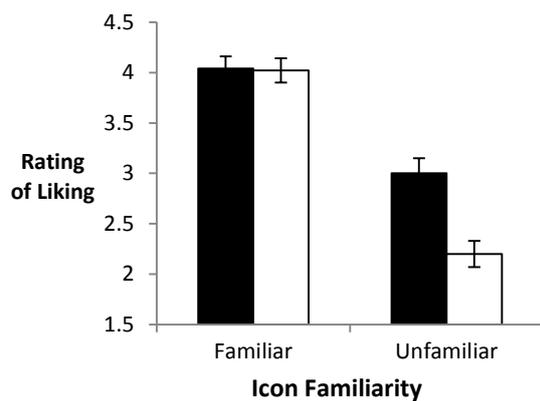
(b) Accuracy Block 9



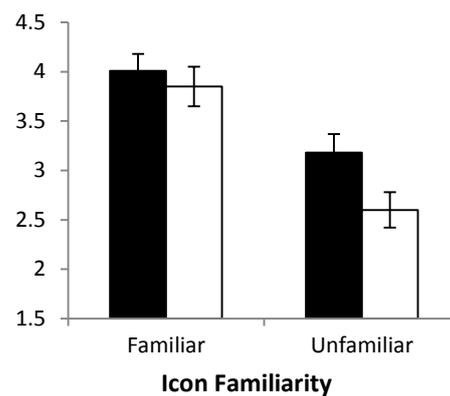
(c) Response Times Block 1



(d) Response Times Block 9



(e) Liking Ratings Block 1



(f) Liking Ratings Block 9

Similarly at Block 9 appeal ratings differed significantly only for unfamiliar icons, $t(20) = 5.87$, $p < .001$, but not familiar icons, $t(20) = 2.22$, $p < .038$. The effects of familiarity and complexity initially (at Block 1) and at the end of the identification task (at Block 9) are illustrated in Figure 6(e) and (f).

As in Experiment 3, the effects of icon familiarity and complexity on performance were reflected in participants' appeal ratings. Importantly, in Experiment 4, where participants were asked to rate appeal at both Block 1 and Block 9, the effects of participants' growing experience with the icon set on performance was also apparent in their ratings. Findings from Experiment 4 are illustrated in Figure 8 (see also Table 3) and can be summarized as follows:-

(i) *Accuracy*

Participants were less accurate and slower at identifying unfamiliar icons at Block 1. By Block 9, participants had learned the icon-function associations and differences in accuracy between familiar and unfamiliar icons were minimal (c.f. Figure 8a and b). Visual complexity had no effect on participants' identification accuracy.

(ii) *Response times*

As in Experiment 3, response times were generally faster for simple icons. Once icon-function associations had been learned, differences in response times remained between simple and complex unfamiliar icons but not for familiar icons (c.f. Figure 8c and d).

(iii) *Appeal*

Mirroring response times, differences in appeal ratings between familiar and unfamiliar icons reduced over time but remained statistically significant even at Block 9. Appeal ratings were lower for unfamiliar complex compared to unfamiliar simple icons.

The aim of Experiment 4 was to examine whether or not learning icon-function relationships could act as a heuristic cue in evaluations of appeal. In Experiment 3 differences between ratings of liking for familiar and unfamiliar icons did not diminish over trials despite the fact that performance measures showed that icon-function relationships were learned. In Experiment 4 appeal ratings were obtained from participants both during initial learning and at the end of the identification task. The assumption was that participants' ratings would reflect the learning that had taken place in between their first and second ratings of the icons (i.e. enhance ease, or fluency, of processing) resulting in an interaction between icon familiarity and blocks of trials, with the effects of familiarity reducing with learning. This was indeed the case. Again, these findings are hard to explain in terms of explicit awareness or instruction since participants were not told that they would gradually learn icon-function relationships as the task progressed or that they should perceive them subjectively to be more appealing as a result.

6 General Discussion

The shift in focus to a more holistic approach in user experience, encompassing users' affective experience as well as usability, has resulted in the visual appeal of interfaces becoming a core construct when considering our affective responses to an interface (Bargas-Avila & Hornbaek, 2011; Brave & Nass, 2008; Thielsch, Blotenberg & Jaron, 2013). Despite

a great deal of research examining interface - particularly website - appeal (e.g., Cyr et al., 2010; Golander et al., 2012; Lindgaard et al., 2011; Moshagen & Theilsch, 2010; Schmidt et al., 2009), very little research has examined the appeal of icons which are an integral part of almost every interface.

Our aim in the experiments reported here was to examine the icon characteristics determining appeal (Experiment 1) and one of the mechanisms which might underpin appeal evaluation, namely the processing fluency heuristic (Experiments 2-4). In particular, we were interested in the possibility that the use of this heuristic might extend beyond the initial instantaneous appeal responses observed for websites and in ERP responses to logos and icons (Handy et al., 2008; Lindgaard et al., 2006, 2011; Thielsch & Hirschfield, 2010)

Experiment 1 established that familiarity and visual complexity, both known to predict efficient performance with icons, are important predictors of icon appeal. The concreteness, or pictorialness, of icons appears to have a secondary role in predicting both performance and appeal. When icons are pictorial they allow us to use our familiarity with what is depicted in order to make inferences about their meaning and so can enhance performance particularly in our initial encounters with them (e.g. Chang & Ng, 2010; Green & Barnard, 1990; McDougall et al., 2000). However, many icons are used for meanings which can be difficult to represent pictorially and we often rely on the frequency with which we have seen an icon and our learning of the icon-meaning relationship in order to understand icons.

The processing fluency heuristic relies on our subjective experiences of the ease or difficulty with which we carry out mental tasks and has been shown to be influential in a wide array of judgments including evaluations of appeal (Alter & Oppenheimer, 2009; Oppenheimer, 2008). This heuristic utilizes the ease with which we carry out mental tasks to make fast and effective judgments by capturing information that 'automatically arrives on our mental stage' (Hertwig et al., 2008, p. 1191). Our working assumption, therefore, was that the ease or difficulty with which we are able to interact with icons on an interface would be reflected, not only in measures of performance (RT, accuracy) but also in our evaluations of their appeal (see Figure 1). If this was indeed the case, we might expect that appeal evaluations would reflect the pattern of observed performance. The visual complexity and familiarity of icons was systematically varied and participants were required to either search for a matching icon in an array (Experiment 2) or to identify an icon given its function (Experiments 3-4).

Performance in both tasks was as predicted on the basis of previous findings (Ben-Bassat & Shinar, 2006; Byrne, 1993, Chan & Ng, 2010; Isherwood & McDougall, 2009; Lesch et al., 2011; McDougall et al., 2000; Shinar et al., 2010). Appeal ratings mirrored performance suggesting that participants were using a processing fluency heuristic when making appeal evaluations (see Figures 4-6). This becomes all the more compelling when one considers that participants could not possibly be consciously aware of the precise pattern of performance in order to reflect this in subjective ratings of liking (since only those with expertise in this area would be likely to be aware of this). Furthermore, appeal ratings for icons were pre-experimentally equated in terms of appeal. This meant that appeal ratings could not be based on initial impressions and depended on the performance differences which emerged as participants gained experience over trials with the icons in the localization and identification tasks.

Appeal evaluations were not always perfectly aligned with performance. In Experiment 3, where participants provided ratings of liking after they had completed *either* 1 *or* 9 blocks of trials, differences in accuracy and response time between familiar and unfamiliar icons reduced over time as participants learned icon-function associations but this was not reflected in appeal ratings. In Experiment 4 we examined the possibility that participants were less aware of learning icon-function associations relative to other available cues by using a repeated measures design in which participants rated icon appeal both in initial and later experimental trials (rather than *either* initial trials *or* later trials as in Experiment 3). When ratings were obtained twice the effects of familiarity on appeal ratings reduced over time, reflecting the learning of icon-function associations. This finding suggests that our implicit awareness of heuristic cues is important in determining their use in appeal evaluations.

This research adds to the considerable body of evidence that processing fluency acts as a cue in determining preferences and appeal across a wide range of tasks (e.g. Alter & Oppenheimer, 2009; Reber et al., 2004). To our knowledge, this is the first time that this has been demonstrated using icons on interfaces and may help to explain the apparent relationship often observed between usability (often assessed via performance) and measures of interface appeal (e.g. Hassenzahl & Monk, 2010). More specifically, this research demonstrates for the first time that the appeal of icons changes dynamically with the pattern of performance: the same set of icons elicited different appeal ratings as a function of task and task demands. Such a dynamic relationship is consistent with recent evidence that performance (or usability) and appeal may operate reciprocally as we interact with an interface to determine user experience (e.g., Reppa & McDougall, 2015; Sauer & Sonderegger, 2014).

The fluency heuristic is likely to be only one item in our inferential toolbox (see Gigerenzer & Goldstein, 1996; Gigerenzer & Gaissmaier, 2011 for a discussion of other heuristics). It may be particularly useful when other information is absent and when quick decisions require frugal automatic, rather than effortful, processing (Hertwig et al., 2008; Marewski & Schooler, 2011). Research to date suggests that a range of other factors are likely to come into play when individuals interact with an interface over time when more considered and conscious decision-making takes place (see Hartmann et al., 2008; Sonderegger & Sauer, 2010; Sonderegger et al., 2012; Thielsch et al., 2013). Thielsch and colleagues (2013) suggest the elaboration likelihood model (Petty & Cacioppo, 1986; Petty & Briñol, 2012, 2015) in which choices are made on the basis of a combination of low-involvement processing (relying on cue-based heuristic processing) and high-involvement processing involving high cognitive processes and consideration of the strength of arguments for and against a particular choice. This approach is akin to a dual-process perspective on evaluative judgments and provides the possibility of an on-going interaction between both types of processing as users experience interfaces (Belke et al., 2015; Dhar & Gorlin, 2013; Kahneman, 2011). The way in which these types of processing interact, particularly over longer periods of interface use, is likely to provide a fruitful area for future research.

7 Conclusion

Although a great deal of research has examined website appeal, remarkably little research has examined icon appeal despite their almost ubiquitous use on interfaces. The experiments reported here examined the processing fluency hypothesis (Alter & Oppenheimer, 2009; Labroo Dhar & Schwarz, 2008; Oppenheimer, 2008; Reber et al., 2004) with regard to icon appeal evaluations. According to this hypothesis, ease of icon processing determines not only user performance but also acts as an implicit heuristic in appeal evaluations. As a result, evaluations of appeal should reflect the patterns observed in task performance.

Experiments 2-4 were rigorous tests of the fluency hypothesis. Icon appeal was made equal prior to experimentation across experimental conditions so that any differences in perceived appeal could therefore only be the result of participants' processing experience during the experiments. In an icon search task (Experiment 2), search times were faster for simpler and more familiar icons and reduced as a result of experience. These effects were mirrored in ratings of appeal. In Experiments 3 and 4 participants were asked to identify icons. Differences in accuracy and response time reduced as participants learned icon-function relationships. Combined Experiments 3 and 4 showed that the predictive power of processing fluency on perceptions of appeal is determined by the availability of processing fluency as a potential heuristic. Taken together the current findings suggest that processing fluency – our experience of ease or difficulty when carrying out mental tasks - is a dynamic mediating factor in determining judgements of appeal.

8 Acknowledgements

The authors gratefully acknowledge the help of Josh Denness, Jessica Emmett and Henrik Waerland in gathering data during the course of these experiments.

9 References

- Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, 36, 140-155.
- Alter, A.L. & Oppenheimer, D.M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality & Social Psychology Review*, 13, 219-235.
- Alter, A.L., Oppenheimer, D.M., Epley, N., Eyre, R.N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology General*; 136, 569–76. doi: 10.1037/0096-3445.136.4.569.
- Bates, E., D’Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., et al. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, 10, 344-380.
- Bargas-Avila, J.A. & Hornbaek, K. (2011). Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. *CHI’ 11 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2689-2698.
- Bauerly, M. & Liu, Y. (2008). Effects of symmetry and number of compositional elements on interface and design aesthetics. *International Journal of Human–Computer Interaction*, 24, 275–287.
- Belke, B., Leder, H., Carbon, C-C (2015). When challenging art gets liked: Evidences for a dual preference formation process for fluent and non-fluent portraits. *PLoS ONE*, 10, e0131796. doi:10.1371/journal.pone.0131796.
- Ben-Bassat, T. & Shinar, D. (2006). Ergonomic guidelines for traffic sign design increase sign comprehension. *Human Factors*, 48, 182-195.
- Berlyne, D.E. (1974). *Studies in the new experimental aesthetics*. Washington, DC: Hemisphere.
- Bornstein, R.F. (1989). Exposure and affect: Overview and meta-analysis of research, 1967-1968. *Psychological Bulletin*, 106, 265-289.
- Brave, S. & Nass, C. (2008). Emotion in human-computer interaction. IN A.Sears & J.A. Jacko (Eds.). *The human-computer interaction handbook: Fundamentals evolving technology and emerging applications, human factors and ergonomics*, pp. 77-92. New York: Taylor and Francis.
- Bundesen, C. (1998). A computation theory of visual attention. *Philosophical Transactions of the Royal Society of London: Series B.*, 353, 1271-1281.
- Byrne, M. D. (1989). Using icons to find documents: Simplicity is critical. Proceedings of INTERCHI ‘93, 446 – 453 (1993) Bornstein, R.F.: Exposure and affect: Overview and meta-analysis of research. *Psychological Bulletin*, 106, 265-289.
- Chan, A.H.S. & Ng, A.W.Y (2010). Effects of sign characteristics and training methods on safety sign training effectiveness. *Ergonomics*, 53, 1325-1346.
- Chi, C.F. & Dewi, R.S. (2014). Matching performance of vehicle icons in graphical and textual formats. *Applied Ergonomics*, 45, 904-916.
- Constable, M.D., Bayliss, A.P., Tipper, S.P., & Kritikos, A. (2013). Self-generated cognitive fluency as an alternative route to preference formation. *Consciousness and Cognition*, 22, 47-52.

- Cyr, D., Head, M., & Larios, H. (2010). Color appeal in website design within and across cultures: a multi-method evaluation. *International Journal of Human-Computer Studies*, 68, 1–21.
- De Wulf, K., Schillewaert, N., Muylle, S., Rangarajan, D. (2006). The role of pleasure in website success. *Information & Management*, 43, 434–446.
- Dhar, R. & Gorlin, M. (2013). A dual-system framework to understand preference construction processes in choice. *Journal of Consumer Psychology*, 23, 528-542.
- Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 24, 515-523.
- Fang, X., Singh, S., & Ahluwalia, R. (2007). An examination of different explanations for the mere exposure effect. *Journal of Consumer Research*, 34, 97-103.
- Forsythe, A., Mulhern, G. & Sawey, M. (2008). Confounds in pictorial sets: The role of complexity in basic level picture processing. *Behaviour Research Methods*, 40, 116-129.
- García, M., Badre, A. N., & Stasko, J. T. (1994). Development and validation of icons varying in their abstractness. *Interacting with Computers*, 6, 191-211.
- Gigerenzer, G. & Goldstein, D.G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review*, 104, 650-669.
- Gigerenzer, G. & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62, 451-82.
- Golander, G.K., Tractinsky, N. & Kabessa-Cohen, I. (2012). Trends in Website Design. *AIS Transactions on Human-Computer Interaction*, 4, 169-189.
- Gordon, P.C. & Holyoak, K.J. (1983). Implicit learning and generalization of the ‘mere exposure’ effect. *Journal of Personality and Social Psychology*, 45, 492-500.
- Green, A. J. K., & Barnard, P. J. (1990). Iconic interfacing: The role of icon distinctiveness and fixed or variable screen locations. In D. Diaper, D. Gilmore, G. Cockton, & B. Shackel (Eds.), *Human computer interaction – Interact '90*, 457–462. Amsterdam: Elsevier Science Publishers.
- Greene, M.R. & Oliva, A. (2009). Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cognitive Psychology*, 58, 137–176.
- Hartmann, J., Sutcliffe, A., & De Angeli, A. (2008). Towards a theory of user judgment of aesthetics and user interface quality. *ACM Transactions on Computer-Human Interaction*, 15, 15–30.
- Handy, T.C., Smilek, D., Geiger, L., Liu, C. & Schooler, J.W. (2008). ERP evidence for rapid hedonic evaluation of logos. *Journal of Cognitive Neuroscience*, 22, 124-138.
- Hassenzahl, M. & Monk, A. (2010). The inference of perceived usability from beauty. *Human-Computer Interaction*, 25, 235-260.
- Hassenzahl, M. & Tractinsky, N. (2006). User experience: A research agenda. *Behavior & Information Technology*, 25, 91-97 .
- Herbert, C., Kissler, J. Junghofer, M., Peyk, P & Rockstroh, B. (2006). Processing emotional adjectives: Evidence from startle EMG and ERPs. *Psychophysiology*, 41, 197-206.
- Herr, P.M. & Page, C.M. (2004). Asymmetric Response Latencies in Liking and Disliking Judgments: So What’s Not to Like? *Journal of Consumer Research*, 30, 588-601.

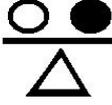
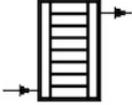
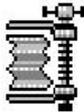
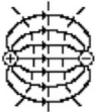
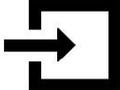
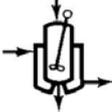
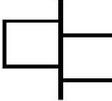
- Hertwig, R., Hertzog, S.M., Schooler, L.J. & Reimer, T. (2008). Fluency heuristic: A model of how the mind exploits a by-product of information retrieval. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 34, 1191-1206.
- Isherwood, S. J., McDougall, S. J. P., & Curry, M. B. (2007). Icon identification in context: The changing role of icon characteristics with user experience. *Human Factors*, 49, 465-476.
- Jacobsen, T. & Höfel, L. (2002). Aesthetic judgments of novel graphic patterns: Analysis of individual judgments. *Perceptual & Motor Skills*, 95, 755-766.
- Joubert, O.R., Rousselet, G.A., Fize, D., Fabre-Thorpe, M. (2007) Processing scene context: Fast categorization and object interference. *Vision Research*, 47, 3286-3297.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Kawabata, H. & Zeki, S. (2004). Neural correlates of beauty. *Journal of Neurophysiology*, 91, 1699-1705.
- Labroo, A.A., Dhar, R & Schwarz, N. (2008). Of frog wines and frowning watches: Semantic priming, perceptual fluency, and brand evaluation. *Journal of Consumer Research*, 34, 819-831.
- Lai, C., Chen, P., Shih, S., Liu, Y. & Hong, J. (2010). Computational models and experimental investigations of the effects of balance and symmetry on the aesthetics of text-overlaid images. *International Journal of Human-Computer Studies*, 60, 269-298.
- Lambon Ralph, M.A., Graham, K.S., Ellis, A.W. & Hodges, J.R. (1998). Naming in semantic dementia – what matters? *Neuropsychologia*, 36, 775-784.
- Lavie, T. & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. *International Journal of Human-Computer Studies*, 60, 269–298.
- Lesch, M.F., Horrey, W.J., Wogalter, M.S. & Powell, W.R. (2011). Age-related differences in warning symbol comprehension and training effectiveness: Effects of familiarity, complexity and comprehensibility. *Ergonomics*, 54, 879-890.
- Lindgaard, G., Fernandes, G., Dudek, C. & Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression! *Behaviour & Information Technology*, 25, 115–126.
- Lindgaard, G., Dudek, C., Devjani, S., Sumegi, L. & Noonan, P. (2011). An exploration of relations between visual appeal, trustworthiness and perceived usability of homepages. *ACM Transactions on Computer-Human Interaction*, 18, doi: 10.1145/1959022.1959023.
- Lui, Y-C. (2005). A simulated study on the effects of information volume on traffic signs, viewing strategies and sign familiarity upon driver's visual search performance. *International Journal of Industrial Ergonomics*, 35, 1147-1158.
- Logan, G. D. (1996). The CODE Theory of Visual Attention: An Integration of Space-Based and Object-Based Attention. *Psychological Review*, 103, 603-649.
- Karvonen, K. (2000). The beauty of simplicity. In: *ACM Proceedings on the 2000 Conference on Universal Usability*. New York: ACM Press, pp. 85–90.
- Mack, Z. & Sharples, S. (2009). The importance of usability in product choice: A mobile phone case study. *Ergonomics*, 52, 1514-1528.

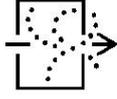
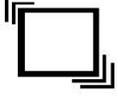
- Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision*, 9, 1-13.
- Manza, L. & Bornstein, R.F. (1995). Affective discrimination and the implicit learning process. *Consciousness and Cognition*, 4, 399-409.
- Manza, L., Reber, A.S. & Zizak, D. (1998). Artificial grammar learning and the mere exposure effect: Emotional preference tasks and the implicit processes. In M.A. Stadler & P.A. Frensch (Eds.), *Handbook of implicit learning*, pp. 201-222. Thousand Oaks, CA: Sage Publications.
- Marewski, J. N., & Schooler, L. J. (2011). Cognitive niches: An ecological model of strategy selection. *Psychological Review*, 118, 3, 393-437.
- McCoy, S., Everard, A., & Lioacono, E.T. (2009). Online ads in familiar and unfamiliar sites: Effects on perceived website quality and intention to reuse. *Information Systems Journal*, 4, 437-458.
- 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.
- McDougall, S. J. P., 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. & Isherwood, S. (2009). What's in a name? The role of graphics, functions and their interrelationship in icon identification. *Behavior Research Methods*, 41, 325-336.
- McDougall, S., Tyrer, V. & Folkard, S. (2006). Searching for signs, symbols, and icons: Effects of time of day, visual complexity, and grouping. *Journal of Experimental Psychology: Applied*, 12, 118-128.
- Monahan, J.L., Murphy, S.T. & Zajonc, R.B. (2000). Subliminal mere exposure: Specific, general, and diffuse effects. *Psychological Science*, 11, 462-466.
- Moshagen, M. & Thielsch, M.T. (2010). Facets of visual aesthetics. *International Journal of Human-Computer Studies*, 68, 689-709.
- Oppenheimer, D. M. (2004). Spontaneous discounting of availability in frequency judgment tasks. *Psychological Science*, 15, 100-105.
- Oppenheimer, D.M. (2008). The secret life of fluency. *Trends in Cognitive Science*, 237-241.
- Orth, U. and Wirtz, J. (2014): Consumer processing of interior service environments: The interplay among visual complexity, processing fluency, and attractiveness. *Journal of Service Research*, 17, 296-309.
- Petty, R. E., & Briñol, P. (2012). The elaboration likelihood model: Three decades of research. In P. A. M. Van Lange, A. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology*, pp. 224–245. London: Sage.
- Petty, R.E. & Briñol, P. (2015). Emotion and persuasion: Cognitive and meta-cognitive processes impact attitudes. *Cognition and Emotion*, 20, 1-26.

- Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, 9, 45-48.
- Reber, R., Schwarz, N. & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, 8, 364-382.
- Reppa, I. & McDougall, S. (2015). When the going gets tough the beautiful get going: Aesthetic appeal facilitates task performance. *Psychonomic Bulletin & Review*, doi:10.3758/s12423-014-0794-z.
- Rogers, Y., & Osborne, D. J. (1987). Pictorial communication of abstract verbs in relation to human-computer interaction. *British Journal of Psychology*, 78, 99-112.
- Rousselet, G.A., Thorpe, S.J., Fabre-Thorpe, M. (2004). Processing of one, two or four natural scenes in humans: the limits of parallelism. *Vision Research*, 44, 877-894.
- Schmidt, K.E., Liu, Y., & Sridharan, S. (2009). Webpage aesthetics, performance and usability: Design variables and their effects. *Ergonomics*, 52, 631-643.
- Scott, D. (1993). Visual search in modern human-computer interfaces. *Behaviour & Information Technology*, 12, 174-189.
- Shinar, D., Dewar, R.E., Summala, H. & Zawkowska, L. (2003). Traffic sign symbol comprehension: A cross-cultural study. *Ergonomics*, 46, 1549-1565.
- Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers*, 28, 516-536.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory*, 6, 174-215.
- Sonderegger, A. & Sauer, J. (2010). The influence of design aesthetics in usability testing: effects on user performance and perceived usability. *Applied Ergonomics*, 41, 403-410.
- Sonderegger, A., Zbinden, G, Uebelbacher, A. & Sauer, J. (2012). The influence of project aesthetics and usability over the course of time: A longitudinal field experiment. *Ergonomics*, 55, 713-730.
- Stotts, D. B. (1998). The usefulness of icons on the computer interface: Effect of graphical abstraction and functional representation on experienced and novice users. In *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, pp. 453-457. Santa Monica, CA: Human Factors Society.
- Theilsch, M.T., Blotenberg, I. & Jaron, F. (2013). User evaluation of websites: From first impression to recommendation. *Interacting with Computers*, 26, 89-102.
- Thielsch, M.T. & Hirschfeld, G. (2010). High and low spatial frequencies in website evaluations. *Ergonomics*, 53, 972-978.
- Tractinsky, N., Katz, A.S. & Ikar, D. (2000). What is beautiful is usable. *Interacting with Computers*, 13, 127-145.
- Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, 12, 97-136.
- Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review*, 95, 15-148.
- Thuring, M. & Mahlke, S. (2007). Usability, aesthetics and emotions in human technology interaction. *International Journal of Psychology*, 42, 253-264.

- van Schaik, P. and Ling, J. (2011). An integrated model of interaction experience for information retrieval in a Web-based encyclopaedia. *Interacting with Computers*, 23, 18-32.
- van Schaik, P. & Ling, J. (2005). Five psychometric scales for online measurement of the quality of human–computer interaction in web sites. *International Journal of Human–Computer Interaction*, 18, 309–322.
- Vartanian, O. & Goel, V. (2004). Neuroanatomical correlates of aesthetic preferences for paintings. *Neuroreport*, 15, 893-897.
- Wang, Q., Cavanagh, P. & Green, M. (1994). Familiarity and pop-out in visual search. *Perception & Psychophysics*, 56, 495-500.
- Wolfe, J.M., Alvarez, G.A., Rosenholtz, R., Kuzmova, Y.I. & Sherman, A.M. (2011). Visual search for arbitrary objects in real scenes. *Attention, Perception & Psychophysics*, 73, 1650-1671.
- Wolfe, J. M., Vo, M. L.-H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and non-selective pathways. *Trends in Cognitive Science*, 15, 77-84.
- Zajonc, R.B. (1968). Attitudinal effects of mere exposure. *Journal of Personality & Social Psychology Monographs*, 9, 1.
- Zajonc, R.B. (1998). Emotions. In Gilbert, Daniel T. (Ed); Fiske, Susan T. (Ed); Lindzey, Gardner (Ed), (1998). *The handbook of social psychology*, (pp. 591-632). New York, NY, US: McGraw-Hill.
- Zajonc, R.B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science*, 10, 224-228.
- Zajonc, R. B. (2000). Feeling and thinking: Closing the debate over the independence of affect. In J. P. Forgas, (Ed.) *Feeling and Thinking: The Role of Affect in Social Cognition*, pp. 31-58.
- Zizek, D.M. & Reber, A.S. (2004). Implicit preferences: The role(s) of familiarity in the structural mere exposure effect. *Consciousness and Cognition*, 13, 336-362.

10 Appendix: Icons used in Experiments 2-4

Familiar complex icons	Familiar simple icons	Unfamiliar complex icons	Unfamiliar simple icons
 Calendar	 Currency exchange	 Bending	 Balance
 Chess	 Diagnose	 Biohazard	 Centre alignment
 Colour	 Educate	 Electrical precipitator	 Eject
 Compress file	 Go rapidly	 Equipotentials	 Electrical loop
 Library	 Speedometer	 Inject resin	 Entrance
 Lift	 Lock	 Iron	 Line vessel
 Note	 Manual control	 Jacketed reactor	 Safety overload device
 Risk of explosion	 Paper industry	 Open half nut	 Spark coil

 <p>Safe</p>	 <p>Picnic area</p>	 <p>Rinse</p>	 <p>Vibrate</p>
 <p>Web crawler</p>	 <p>Slow</p>	 <p>Steriliser</p>	 <p>Zoom</p>