

What's in a name? The role of graphics, functions, and their interrelationships in icon identification

SINÉ MCDUGALL

Bournemouth University, Bournemouth, England

AND

SARAH ISHERWOOD

University of Leeds, Leeds, England

Communication using icons is now commonplace. It is therefore important to understand the processes involved in icon comprehension and the stimulus cues that individuals utilize to facilitate identification. In this study, we examined predictors of icon identification as participants gained experience with icons over a series of learning trials. A dynamic pattern of findings emerged in which the primary predictors of identification changed as learning progressed. In early learning trials, semantic distance (the closeness of the relationship between icon and function) was the best predictor of performance, accounting for up to 55% of the variance observed, whereas familiarity with the function was more important in later trials. Other stimulus characteristics, such as our familiarity with the graphic in the icon and its concreteness, were also found to be important for icon design. The theoretical implications of these findings are discussed, with particular emphasis on the parallels with picture naming. The icon identification norms from this study may be downloaded from brm.psychonomic-journals.org/content/supplemental.

Johnson, Paivio, and Clark (1996) noted that “picture naming is . . . of practical importance because pictures convey important information, either alone or in conjunction with written text. Overt or covert naming, therefore, occurs in many educational, occupational, and recreational activities” (p. 113). This statement is particularly true of icons, since their major function is to communicate information on computers, in public spaces and roads, in transport (e.g., aircraft and cars), and on home equipment (e.g., telephones, washing machines, and video recorders). It is important to understand more about the way in which icons are identified, because they are encountered and used at least as frequently as pictures in today's world. For the sake of simplicity, *icons* is the term used in this article to refer to the broad range of icons, signs, or symbols used to help individuals interact with machines and their environment.

Icons differ in a number of ways from the pictures typically employed in naming tasks (Cycowicz, Friedman, Rothstein, & Snodgrass, 1997; Snodgrass & Vanderwart, 1980; cf. Hancock, Rogers, Schroeder, & Fisk, 2004, and Isherwood, McDougall, & Curry, 2007). One difference is that icons are not always objects and may be quite abstract representations of the functions that they represent. For example, an icon indicating a lift (an elevator) may depict individuals in a lift (see Figure 1A) or may simply use up and down arrows to refer to the motion of the lift

(see Figure 1B). Another difference is that strong semantic links between pictures and their names have usually been forged over a long period, whereas the intended function of an icon often has to be learned. The amount of learning required is likely to depend on the nature of the relationship between the icon and its function (for a discussion of these relationships, see Familant & Detweiler, 1993; Keller & Stevens, 2004; Peirce, 1932). Figure 1 illustrates the range of these relationships. In some cases, simply naming the graphic in the icon is sufficient to arrive at the function name (naming the lock in Figure 1C). In other cases, the icon–function relationship is less direct and must be inferred (e.g., inferring that hares run quickly, as in Figure 1D). In other cases, the relationship is arbitrary and can be understood only if users have previously learned its meaning (the relatively well-known shape indicating biohazards in Figure 1E).

Finally, the range of functions that icons are designed to communicate are broader than those usually considered in the picture-naming literature, and again this is illustrated in Figure 1. In order to predict the ease with which icons are likely to be identified, the icon characteristics of each part of the icon need to be considered: (1) the picture, or graphic, used in the icon; (2) the function of the icon; and (3) the relationship between icon and function. This is similar to the three classes of item attributes expected to influence picture naming (Johnson et al., 1996; Paivio, Clark,

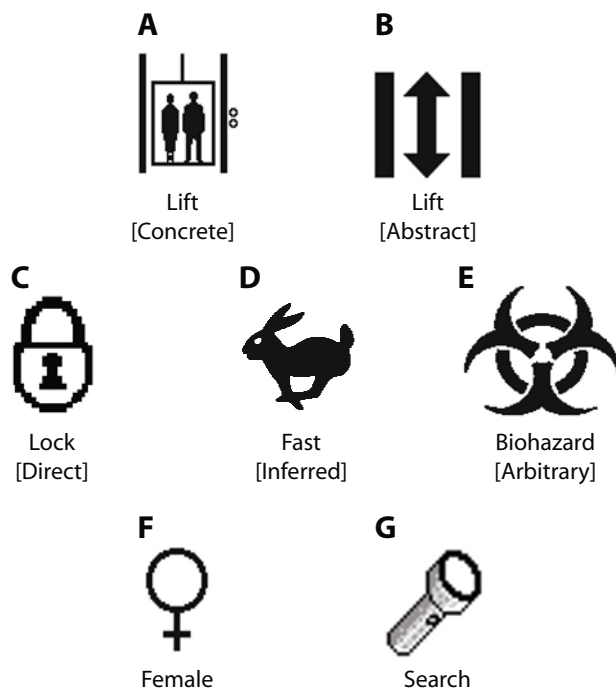


Figure 1. Examples of different types of icons.

Digdon, & Bons, 1989). Measures of these attributes are typically obtained in picture-naming studies and include (1) visual complexity and familiarity of the picture; (2) the frequency of occurrence, familiarity, age of acquisition, word length, phonological complexity, imageability, and concreteness of the name; and (3) the closeness of the relationship between picture and name, as assessed using measures of goodness of depiction, name agreement, and concept agreement (e.g., Alario et al., 2004; Barry, Morrison, & Ellis, 1997; Bates et al., 2003; Bogka et al., 2003; Bonin, Barry, Méot, & Chalard, 2004; Bonin, Chalard, Méot, & Fayol, 2002; Catling & Johnston, 2006; Cuetos, Ellis, & Alvarez, 1999; Dell'Acqua, Lotto, & Job, 2000; Ellis & Morrison, 1998; Lloyd-Jones & Nettlemill, 2007; Morrison & Gibbons, 2006; Morrison, Hirsh, & Duggan, 2003; Snodgrass & Yuditsky, 1996; Weekes, Shu, Hao, Liu, & Tan, 2007; Zevin & Seidenberg, 2002).

Although the role of the characteristics of the icon graphic and the closeness of the icon–function relationship have been investigated in icon identification, research examining the role of function characteristics has been more limited and has been concerned primarily with the inclusion or exclusion of text labels to accompany icons. This research has shown that individuals often rely on textual labels attached to icons to access meaning (Guastello, Traut, & Korienek, 1989), particularly where the icon graphic is unfamiliar (Vukelich & Whitaker, 1993), and suggests that the characteristics of the icon function label are likely to be important in determining user performance with an interface. The aim of the present research was to find out more about how the nature, or characteristics, of icon function names affects the ease with which icons are identified and to ascertain the importance of these charac-

teristics relative to measures of icon graphic characteristics and the closeness of the icon–function relationship.

Because icons are often not known and have to be learned, predictors of icon identification change as learning occurs (Green & Barnard, 1990; Isherwood et al., 2007; McDougall, de Bruijn, & Curry, 2000). Changes as a result of learning were therefore also examined by obtaining accuracy and response time (RT) data as participants gained experience with an icon set over a series of 13 blocks of learning trials. Again, this contrasts with picture-naming experiments where, because item names are known, participants are usually given one, or at most two, exposures to stimuli.

Icon Characteristics

Visual complexity. The visual complexity of an icon or picture can be measured in a number of ways (see Forsythe, Sheehy, & Sawey, 2003; García, Badre, & Stasko, 1994; Lloyd-Jones & Luckhurst, 2002; Snodgrass & Vanderwart, 1980) but generally can be taken to refer to the level of intricacy or detail in a stimulus. Complexity metrics typically assess a combination of the number of lines and shapes out of which the stimulus is composed. Some studies have shown that visual complexity increases picture-naming times, and this is thought to be the result of increased processing time at, or before, the stage of object recognition (Alario et al., 2004; Ellis & Morrison, 1998; Humphreys, Riddoch, & Quinlan, 1988). The complexity effect, however, is far from robust (Barry et al., 1997; Bonin et al., 2002; Bonin, Peereman, Malardier, Méot, & Chalard, 2003; Cuetos et al., 1999; Paivio et al., 1989; Snodgrass & Yuditsky, 1996), and some researchers have even suggested that increased complexity can

enhance performance (Biederman, 1987; Lloyd-Jones & Luckhurst, 2002). One possible reason for the lack of robustness is that subjective ratings of visual complexity are confounded with familiarity, so that more familiar items are rated as less complex (Forsythe, Mulhern, & Sawey, 2008; Székely & Bates, 2000). Any effects of complexity on icon identification may, therefore, be suppressed in regression analyses by familiarity. Another reason, apparent from icon research, is that visual complexity is important when we need to *search* visual displays (Byrne, 1993; McDougall et al., 2000, Experiment 1; McDougall, Tyrer, & Folkard, 2006) but has less of a role to play when icons must be *identified* (McDougall et al., 2000, Experiment 2). It is more difficult to explain the opposing finding, that increased complexity enhances performance, unless there is a dual locus for complexity effects, one associated with visual search (where reduced complexity is helpful) and another associated with identification (where increased complexity is helpful; see Lloyd-Jones & Nettlemill, 2007). If there is a dual locus, visual complexity is likely to affect the identification task. If, on the other hand, complexity is associated primarily with visual search, the effects of complexity on icon identification will be small, and where they do arise, they are likely to be confounded with icon familiarity.

Icon familiarity. In the picture-naming literature, rated familiarity is a strong and consistent predictor of picture naming (Alario et al., 2004; Bates et al., 2003; Ellis & Morrison, 1998; Feyereisen, Van der Borgh, & Seron, 1988; Hirsh & Funnell, 1995; Snodgrass & Yuditsky, 1996). Familiarity is thought to be the equivalent of frequency for words (Bates et al., 2003; Hirsh & Funnell, 1995) and to reflect the ease with which semantic representations can be accessed in long-term memory (Hirsh & Funnell, 1995; Lambon Ralph, Graham, Ellis, & Hodges, 1998).

Recent research has shown that familiarity with the graphic in the icon is an important predictor of ease of identification even after individuals have gained experience with the icons over a series of trials (Isherwood et al., 2007). This was also thought to reflect differences in ease of access to long-term representations. Research to date therefore suggests that icon familiarity would be an important predictor of identification across the whole series of learning trials.

Icon concreteness. The concreteness of an icon is the extent to which it depicts real objects, materials, or people (see Figures 1A and 1B). Abstract icons, in contrast, are less pictorial and tend to have less obvious connections with real-world items and to make more use of shapes, arrows, and lines (McDougall, Curry, & de Bruijn, 1999). Initial research suggested that concreteness was the most important property of an icon because concrete icons depict objects, allowing individuals to use their knowledge of the everyday world to access meaning (Moyes & Jordan, 1993). This was supported by research showing that users respond more quickly and accurately to concrete icons (Arend, Muthig, & Wandmacher, 1987; Rogers & Osborne, 1987; Stammers, George, & Carey, 1989; Stammers & Hoffman, 1991; Stotts, 1998). Other research, however, has shown that these performance advantages

are not long lasting and diminish as users gain experience with icons (Green & Barnard, 1990; Isherwood et al., 2007; McDougall et al., 2000). It was therefore expected that the locus of concreteness effects would be in initial learning trials as individuals attempt to form meaningful links between icon graphics and their function. Concreteness effects were also expected to be limited by the extent to which the pictorialness of the icon allowed access to *appropriate* function names via visual metaphors (e.g., simply naming the rabbit in Figure 1D does not allow access to the appropriate function name).

The Icon–Function Relationship

Measures of the strength of item–referent relationships are important predictors of naming performance. Bates et al. (2003) used a measure of goodness of depiction for pictures where participants were asked to judge how well each picture illustrated its given name, using a 1–5 scale. Other associated measures have been obtained for both pictures and icons. These include concept agreement (the percentage of participants correctly naming the item) and, more commonly, name agreement (which assesses the number of names given to an item across participants; Cycowicz et al., 1997; Fiez & Tranel, 1997; Lachman & Lachman, 1980; Martein, 1995; McDougall et al., 1999; Morrison et al., 2003; Sanfeliu & Fernandez, 1996; Snodgrass & Vanderwart, 1980; Snodgrass & Yuditsky, 1996). Bates et al.'s measure of goodness of depiction was the most important predictor of naming latency across seven languages, and name agreement is also a robust predictor of naming performance (e.g., Alario et al., 2004; Barry et al., 1997; Johnson & Clark, 1988; Morrison et al., 2003; Snodgrass & Yuditsky, 1996; Vitkovitch & Tyrrell, 1995).

Peirce's (1932) original tripartite classification system (see Figures 1C–1E) was unified into a single continuous dimension, semantic distance, by McDougall et al. (1999) when they obtained ratings of the strength of the icon–function relationship on a 1–5 scale from *very strongly related* to *not closely related*, providing a measure similar to the goodness-of-depiction measure employed by Bates et al. (2003). When McDougall, Curry, and de Bruijn (2001) varied both the semantic distance and the concreteness of an icon set, they found that semantic distance was a stronger determinant of performance than was concreteness (see also Isherwood et al., 2007). This may well be because concreteness acts as an index of how well we name pictured items but is not an index of whether or not we can access the appropriate function meaning. In the present study, semantic distance was therefore expected to be a strong predictor of naming, acting in a way similar to that of Bates et al.'s measure.

Function Characteristics

Three function characteristics were considered in this study: familiarity, concreteness, and function name length. The recent and continuing development of icons as a form of communication means that measures of age of acquisition and frequency of occurrence typically used in picture naming are not yet available for icons.

Familiarity. It is impossible to consider name familiarity without considering name frequency, since the two have been closely linked in research. In the first studies showing a *frequency* effect on naming latency, Oldfield and Wingfield (1964, 1965) argued that naming times for pictures should be related to the *familiarity* of the object and its name. In a later review of the picture-naming literature, Johnson et al. (1996) also claimed that word frequency may index the familiarity, and semantic accessibility, of the concept underlying that word. This idea is supported by studies that have shown that frequency effects are not apparent unless semantic access is required (Caramazza, Costa, Miozzo, & Bi, 2001; Meyer, Sleiderink, & Levelt, 1998). If familiarity with the function name is implicated in semantic access (as is familiarity with the graphic), name we would expect that familiarity would be a strong predictor of naming even when the icon set has been learned.

Function name length. Word length is most often measured using the number of syllables, letters, or phonemes that a word contains. Word length effects are typically associated with the word production, or pronunciation, stage of naming (Johnson et al., 1996; Levelt, 1989; Levelt, Roelofs, & Meyer, 1999). However, word length effects in picture naming are often small and inconsistent (see Cuetos et al., 1999; Dell'Acqua et al., 2000; Santiago, MacKay, Palma, & Rho, 2000; Snodgrass & Yuditsky, 1996). A possible reason for this is the negative correlation between frequency and word length (i.e., Zipf's [1965] law). This relationship occurs because short words tend to be used for more frequently occurring concepts and longer names for less frequently occurring concepts. This means that, across languages, negative correlations are observed between frequency and word length (see Bates et al., 2003). Word length effects therefore may be suppressed in regression analyses because of correlations between the frequency with which we have encountered a concept (i.e., our familiarity with the function) and the length of the function name.

Function concreteness. Word imageability (the ease with which a mental image can be formed to a word) is closely correlated with word concreteness (the extent to which words refer to objects or people; Gilhooly & Logie, 1980; Paivio, Yuille, & Madigan, 1968; $r_s = .83$ and $.78$, respectively). Despite these high correlations, Bird, Franklin, and Howard (2001) point out that imageability is a better predictor of picture-naming performance than is concreteness (see also Nickels & Howard, 1995). This may be because the pictures, by their very nature, are highly concrete items (they are objects or people) and this produces ceiling effects that limit concreteness's predictive power. Given that the concreteness of icon functions varies considerably, it may well be that concreteness is a better predictor of icon naming.

It is unlikely that function and icon concreteness will act similarly as predictors of icon identification, because word and icon concreteness are often dissociated. Figures 1F and 1G illustrate this dissociation; *female*, a word rated as highly concrete, is depicted using an abstract icon,

whereas *search*, a relatively abstract word, is depicted using a concrete icon. Exactly what the role of function concreteness would be with respect to icon naming was difficult to predict, given the lack of relevant previous research, so no specific hypotheses were formed prior to experimentation. Ratings of function name concreteness were obtained prior to conducting the icon identification trials, since no measures were currently available.

Icons are used today as frequently as pictures, so it is important to ascertain the factors determining ease of icon identification so that icon design can be improved. To date, no research has examined the role of the function name in determining identification. One of the reasons for this may be a practical one: The function is usually a given, and it is the interface designer's task to create an icon that works well for that function. In the research reported here, we examined the relative predictive power of characteristics associated with the function name, the icon graphic, and the strength of the icon–function relationship in determining ease of identification as participants learned icons and their functions during a series of trials. In Experiment 1, additional ratings of function name characteristics were obtained, and in Experiment 2, predictors of icon identification were examined.

EXPERIMENT 1 Additional Ratings of Function Name Characteristics

Ratings of icon characteristics (concreteness, familiarity, and complexity) and icon–function relationships (semantic distance, concept agreement, and name agreement) have already been obtained for a corpus of icons by McDougall et al. (1999). However, there has been no previous investigation of the role of function names on icon identification, so no measures of function concreteness or familiarity currently exist. These were obtained in Experiment 1 by extending the ratings obtained by McDougall et al. (1999) to include function concreteness and familiarity.

Method

Participants. Eighty-three participants provided ratings for function name characteristics. All were from an introductory psychology class at the University of Wales, Swansea, and received course credit for their participation. Forty-three provided ratings of function name concreteness (mean age = 20 years 10 months; $SD = 5$ years 10 months), and 40 provided ratings of function name familiarity (mean age = 20 years; $SD = 4$ years). Four participants' data were excluded because the ratings they provided were incomplete (1 providing familiarity ratings and 3 providing concreteness ratings).

Materials and Procedure. The procedure used to obtain these ratings was the same as that employed by McDougall et al. (1999) when obtaining ratings for other icon characteristics. Function names were presented to the participants in booklets. Pages in booklets were assembled in accordance with a Latin square design to ensure that each participant was presented with the icon functions in a different order. Alongside each function name was a 5-point rating scale. Following the procedures developed by Paivio et al. (1968) and Gilhooly and Logie (1980), the participants providing concreteness ratings were told that function names were to be regarded as concrete if they depicted real objects, materials, or people; those that did not were to be regarded as abstract. A rating of 5 indicated

that the function name was *definitely concrete*, and a rating of 1 indicated that it was *definitely abstract*. Familiarity ratings also used a 1–5 scale; a rating of 5 indicated that the function was *very familiar*, and a rating of 1 indicated that it was *very unfamiliar*.

Results and Discussion

A downloadable appendix to this article provides an alphabetical index of the concreteness and familiarity of function name ratings and number of syllables in each function. Icons can be matched with those in the McDougall et al. (1999) study, which also lists icons alphabetically and with the same numbers. One item, *belt drive*, was omitted in error from the ratings booklets, leaving a total of 238 icons from the original icon set. There was a considerable spread in ratings for both the concreteness and familiarity of function names obtained (function familiarity, $M = 3.62$, $SD = 1.13$, range = 1.18–5.00; function concreteness, $M = 3.02$, $SD = 0.76$, range = 1.85–4.68; number of syllables, $M = 3.81$, $SD = 2.40$, range = 1–13). Although there was a degree of skew for all three variables (function familiarity = $-.64$, function concreteness = $.71$, number of syllables = 1.00), it was well within acceptable limits (George & Mallery, 2005, p. 99), and no transformations were carried out.

The reliability of the ratings was evaluated using split-half reliability. The participants were split randomly into two groups; this resulted in two groups of 19 and 20 participants for each rating. Correlations between subgroup ratings were then calculated. These were high for both ratings ($r = .90$ for concreteness and $r = .97$ for familiarity). Another indication of reliability is available, since a number of icons from the original corpus had the same function name (as in *lift* in Figure 1). Since the function name alone was presented, both concreteness and familiarity ratings should be very similar, and this was indeed the case (see *debug*, *films*, *lift*, *no entry*, *search*, and *zoom* in the downloadable appendix).

Interrelationships between item characteristics.

The correlations between the icon characteristics, icon–function relationships, and function name characteristics are shown in Table 1. Correlations for icon characteristics and icon–function relationships are taken from McDougall et al. (1999) and will not be considered in detail here. In

general, there are strong correlations between these characteristics, with the exception of the visual complexity ratings.

Function name concreteness correlates relatively poorly with icon concreteness and other icon characteristics. Possible reasons for this become apparent if one examines *lift* (in Figure 1B), *fast*, *biohazard*, *female*, and *search* in Figure 1, where there is a dissociation between the concreteness of the graphic and the name. Function concreteness correlates relatively well with icon–function relationship measures, reflecting the fact that concrete functions can be more easily represented visually in an icon, creating a close icon–function relationship.

With the exception of visual complexity, there are strong correlations between function name familiarity and both icon characteristics and icon–function indices. These correlations suggest that our familiarity with the function name may, in combination with the icon familiarity, help to determine ease of access to function meaning. The strongest correlation with function name length (number of syllables) is with function name familiarity, as might be expected on the basis of Zipf’s (1965) law.

EXPERIMENT 2

The Role of Function Name Characteristics in Predicting Icon Identification

In this experiment, we examined the role of function name characteristics in icon identification as participants gained experience with the icon set over a series of trials. Of particular interest was the role of function name characteristics, relative to icon graphic and icon–function characteristics, in determining identification.

Method

Participants. Thirty undergraduate and postgraduate volunteers from the University of Wales, Swansea, took part in this experiment. They were either paid or given psychology course credits for their participation. One participant’s data were rejected because the participant did not follow the experimental procedure correctly, leaving a total of 29 participants who completed the icon identification task.

Materials and Apparatus. The icons used in this experiment were chosen from the set of 239 icons from the original corpus. One hundred eight icons were selected for use in the experimental task;

Table 1
Correlations Between Icon Characteristics From the McDougall, Curry, and de Bruijn (1999) Icon Set and Function Name Characteristics Obtained in Experiment 1 ($n = 238$)

Characteristic	Icon Concreteness	Icon Familiarity	Visual Complexity	Semantic Distance	Name Agreement	Function Concreteness	Function Familiarity	No. of Syllables
Icon								
Concreteness	–							
Familiarity	.76**	–						
Visual complexity	–.06	–.34**	–					
Icon–Function Relationship								
Semantic distance	.65**	.53**	–.05	–				
Name agreement	.73**	.74**	–.23*	.61**	–			
Function Name								
Concreteness	.18*	.14	–.06	.41**	.23**	–		
Familiarity	.51**	.46**	–.06	.51**	.44**	.38**	–	
No. of syllables	–.27**	–.23**	.05	–.36**	–.28**	–.33**	–.61**	–

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

Table 2
Means and Standard Deviations for the Characteristics
of the 72 Icons Employed in Experiment 2

Characteristic	<i>M</i>	<i>SD</i>	Range
Icon			
Concreteness	3.12	1.01	1.65–4.93
Familiarity	2.92	0.97	1.53–4.87
Visual complexity	2.63	0.87	1.04–4.23
Icon–Function Relationship			
Semantic distance	2.48	0.92	1.17–4.88
Name agreement (%)	30.08	24.20	2.50–87.50
Function Name			
Concreteness	3.03	0.73	1.98–4.68
Familiarity	3.57	1.13	1.18–5.00
No. of syllables	4.04	2.45	1–10

36 of the icons were used in practice trials, and the remaining 72 icons were used in experimental trials. The initial selection of the 72 experimental icons was made randomly, and then some adjustments were made to ensure that the full range of characteristics under examination was represented. The means, standard deviations, and range of the icon characteristics of this icon set are shown in Table 2.

A PC equipped with an Intel Pentium 166-MHz processor was used to present experimental trials in the icon identification task. The participants used a mouse to provide input for the program, with the track speed set midway between slow and fast. Responses were timed using the system's multimedia timer, which allows measurement to within 1 msec, although an estimated uncertainty of up to ± 30 msec is added when the mouse is used as an input device.

Procedure. In this task, the participants were shown a function label and were required to select the icon that they felt best matched this function from an array of nine icons (see Figure 2). This task is analogous to the kind of task users face when interacting with interfaces when they need to find icons that match something that they wish to do (see Bocker, 1993, p. 76, for a discussion of this type of task). The procedure for each experimental trial was as follows.

1. The participants clicked on the "OK" button presented on the left-hand side of the screen. The function of an icon would then appear on the screen.

2. After 2 sec, a bleep sounded, indicating that the participants should click the "OK" button again. Given that it takes an average of approximately 500 msec for a skilled adult reader to read and pronounce a single word, it was expected that the participants would be able to read function words or phrases easily within this 2-sec period (Waters, Seidenberg, & Bruck, 1984). Previous research and extensive piloting also suggested that this time was sufficient for the participants (see McDougall et al., 2000).

3. The array of nine icons then appeared on screen, and the participants selected, as quickly as possible, the icon they felt matched the function label.

4. After the participants had selected an icon, the icons disappeared, and the screen returned to its original state, ready for the participants to begin the next experimental trial.

This procedure was repeated until each of the 72 icons had been displayed as a target icon. For each experimental trial, timing began when the 9-icon array appeared on the screen and ceased after the participants had clicked on 1 of the icons. Whenever the participants did not select the correct icon, they were given two more attempts to do so. After their third incorrect attempt, all of the other icons in the display disappeared, leaving only the target icon visible for 2 sec before it also disappeared, after which the next trial began.

Each of the icons used in this experiment was presented once as a target icon and eight times as a distractor icon in each block of 72 trials. Because all of the icons were shown equally often, both as targets and as distractors, this controlled for possible differences in the incidental learning of icons. The location of the distractors was randomized across each block of trials. Prior to experimental trials, the participants were presented with 36 practice trials, using non-experimental icons, in order to familiarize them with the experimental procedure. This was followed by 13 blocks of 72 experimental trials, which were presented to the participants in two sessions on consecutive days. The participants were given a 2-min break between the blocks of learning trials. In total, the participants were shown each icon 13 times as a target and 104 times as a distractor.

Results and Discussion

Mean accuracy scores and RTs for each block of experimental trials are shown in Table 3. Also shown are the correlations between item characteristics and performance measures. The pattern of correlations differed for accuracy scores and RTs. Both icon concreteness and familiarity were significantly correlated with accuracy scores, although this was reduced as participants gained experience with the icons over blocks of trials. Semantic distance was strongly correlated with accuracy in early blocks of trials. Name agreement was correlated significantly with accuracy across all the blocks of trials; this may reflect its close relationship, noted earlier, with both icon concreteness and familiarity. Both function familiarity and function length (number of syllables) were significantly related to accuracy, with familiarity showing the most consistently high correlations with performance across all the blocks of trials, when compared with all other dependent vari-



Figure 2. An experimental trial.

Table 3
Mean Accuracy and Response Times (With Standard Deviations) and
Correlations Between These Performance Measures and Icon Characteristics for Each Block of Trials

	Block												
	1	2	3	4	5	6	7	8	9	10	11	12	13
	Accuracy (% Correct)												
<i>M</i>	70.16	89.42	95.40	96.74	97.80	97.84	98.04	98.56	98.71	98.61	98.99	98.90	98.80
<i>SD</i>	21.71	13.44	7.25	3.59	3.44	4.11	2.82	2.94	2.34	2.75	1.87	2.37	2.47
Icon													
Concreteness	.46**	.52**	.42**	.40**	.30**	.39**	.30**	.28*	.23	.25*	.10	.22	.29*
Familiarity	.47**	.50**	.41**	.35**	.29*	.35**	.30**	.34**	.23	.29*	.26*	.25*	.28*
Visual complexity	-.04	-.01	.11	-.11	.03	-.01	-.06	-.18	-.17	-.10	-.24	-.19	-.18
Icon-Function Relationship													
Semantic distance	.74**	.63**	.42**	.46**	.31**	.44**	.28*	.23	.18	.21	.20	.18	.09
Name agreement	.44**	.44**	.40**	.36**	.31**	.35**	.28*	.29*	.14*	.27*	.24*	.20*	.26*
Function Name													
Concreteness	.18	.08	.11	.13	.03	.14	.30**	.13	.19	-.07	-.01	.17	.02
Familiarity	.62**	.65**	.59**	.47**	.46**	.58**	.54**	.44**	.40**	.47**	.15	.38**	.43**
No. of syllables	-.28*	.36**	-.51**	-.27*	-.38**	-.40**	-.49**	-.45**	-.09	-.36**	-.09	-.24*	-.26*
	Response Times (msec)												
<i>M</i>	2,924	1,948	1,666	1,577	1,485	1,416	1,361	1,340	1,321	1,267	1,272	1,225	1,249
<i>SD</i>	881	636	447	361	344	299	254	234	216	196	202	183	194
Icon													
Concreteness	-.36**	-.55**	-.56**	-.55**	-.43**	-.35**	-.37**	-.31**	-.49**	-.42**	-.40**	-.32**	-.34**
Familiarity	-.44**	-.59**	-.61**	-.54**	-.45**	-.36**	-.43**	-.47**	-.52**	-.52**	-.43**	-.41**	-.43**
Visual complexity	.13	.13	.24*	.36**	.22	.18	.30**	.34**	.33**	.34**	.35**	.38	.38**
Icon-Function Relationship													
Semantic distance	-.74**	-.71**	-.63**	-.53**	-.56**	-.50**	-.46**	-.46**	-.52**	-.47**	-.46**	-.41**	-.35**
Name agreement	-.35*	-.51**	-.55**	-.55**	-.45**	-.37**	-.41**	-.36**	-.48**	-.46**	-.44**	-.39**	-.40**
Function Name													
Concreteness	-.18	-.11	-.19	-.19	-.18	-.23	-.20	-.22	-.24*	-.24*	-.25*	-.31**	-.32**
Familiarity	-.33**	-.52**	-.60**	-.52**	-.55**	-.56**	-.58**	-.52**	-.50**	-.51**	-.53**	-.44**	-.42**
No. of syllables	.10	.28*	.38**	.39**	.50**	.49**	.50**	.37**	.40**	.47**	.40**	.42**	.40**

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

ables. Neither visual complexity nor function concreteness (with the exception of Block 7) was related to accuracy of identification; they were therefore not included in the subsequent analyses of accuracy scores.

Correlations between icon characteristics and RTs were generally higher, particularly in later blocks of trials. In the initial blocks of trials, the strongest correlations were with semantic distance. Both icon familiarity and function familiarity were also strongly correlated with identification time across all the blocks of trials. It is interesting to note that both function concreteness and visual complexity were correlated significantly with RT in later blocks of trials. Correlations after seven blocks of trials, however, should be treated with considerable caution, particularly for accuracy scores, since performance appeared to have reached asymptote and variation was therefore limited. For this reason, only Blocks 1-7 will be considered in the analyses that follow.

Prior to ANCOVAs and regressions being carried out, consideration was given to the risk of multicollinearity between icon characteristics. Tabachnick and Fidell (1996) recommended that including variables that correlate above .70 with one another should be avoided (see also Morrison, 2003, for further discussion of the effects of multicollinearity in picture-naming analyses). Table 1 shows that icon familiarity was correlated highly with both icon concreteness (*r* = .76) and name agreement (*r* = .73). Recent research (Isherwood et al., 2007) and the correlations with

performance shown in Table 3 suggest that icon familiarity is likely to be the strongest predictor of the three variables for both accuracy and RTs. Furthermore, preliminary regression analyses revealed that name agreement and icon concreteness predicted little, if any, variance in accuracy or RTs in any of the blocks of trials after icon familiarity had been considered. As a result, these two variables were excluded from further analyses.

Previous research has shown that word length effects are often minimal because of Zipf's (1965) law (see Bates et al., 2003). Preliminary regression analyses were therefore carried out for word familiarity and the number of syllables in the function name. These revealed that the length of the function name predicted little variance in performance measures across all the blocks of trials once function familiarity had been considered; number of syllables was therefore also excluded from further analyses.

Accuracy. A repeated measures ANCOVA was conducted in order to examine the effect of learning across blocks of trials (1-7) on accuracy scores. Covariates were icon familiarity, semantic distance, and function name familiarity. These were the characteristics remaining after lack of correlation with performance (visual complexity and function concreteness) and collinearity with other variables (icon concreteness, name agreement, and function length) led to the exclusion of a number of variables. A by-items, rather than by-subjects, analysis was carried out, since it was the nature of the icons that was the focus of the study.

There were significant increases in accuracy as the participants learned the icon set across blocks of trials [$F(6,408) = 123.16, MS_e = 0.440, p < .01, \eta_p^2 = .64$]. Two of the three covariates, semantic distance and function name familiarity, had significant overall effects on accuracy [$F(1,68) = 22.29, MS_e = 0.280, p < .01, \eta_p^2 = .25$, and $F(1,68) = 22.80, MS_e = 0.286, p < .01, \eta_p^2 = .25$, respectively], but icon familiarity did not [$F(1,68) = 2.43, MS_e = 0.03, p > .05$]. Similarly, interactions with blocks of trials were significant for semantic distance [$F(6,408) = 25.86, MS_e = 0.092, p < .01, \eta_p^2 = .28$] and function name familiarity [$F(6,408) = 5.53, MS_e = 0.020, p < .01, \eta_p^2 = .08$], but not for icon familiarity [$F(6,408) = 1.13, MS_e = 0.004, p > .05$].

The relationship between the significant covariates and accuracy was examined further, using regression analyses for each block of trials. These analyses had three aims as follows: (1) to examine which characteristic was the best predictor of accuracy, which information is important on a practical level since, if one characteristic is a sufficiently strong predictor, it can be used as an index of icon usability; (2) to examine the extent to which each characteristic predicted unique variance in accuracy scores across trials and the extent to which the characteristics overlapped in the variance they explained; and (3) to examine the combined predictive power of the variables considered.

Table 4 shows the results of these analyses: the percentage of variance explained when each covariate was entered first into the regression equation (the extent to which each predicted accuracy scores), the percentage of unique variance predicted after the other covariates had been considered, and the total percentage of the variance in accuracy scores predicted by the combination of all the variables considered. Semantic distance was the strongest predictor of accuracy in the first block of trials as the par-

ticipants tried to learn the icon–function relationships, but from Block 3 onward, function name familiarity was the primary determinant of performance, and it contributed uniquely to explaining variance in accuracy across all the blocks of trials. Together, these two icon characteristics explained over 60% of the variance seen in accuracy in the first block of trials. Not surprisingly, this was reduced as the icon set was learned, but even when the learning process was slowing in Blocks 6 and 7, icon familiarity accounted for over 30% of the variance observed.

Response times. The same type of analyses were carried out for the RT data. A repeated measures ANCOVA was used to consider the effect of learning on performance across blocks of trials. In addition to icon familiarity, semantic distance, and function name familiarity, visual complexity and function name concreteness were included as covariates, since they were correlated significantly with RTs in later blocks of trials.

RTs were reduced considerably across the learning curve [$F(6,396) = 18.24, MS_e = 1,331,750.92, p < .01, \eta_p^2 = .21$]. As with the accuracy scores, semantic distance and function name familiarity had significant overall effects on RTs [$F(1,66) = 38.77, MS_e = 15,618,784.45, p < .01, \eta_p^2 = .37$, and $F(1,66) = 4.38, MS_e = 1,765,289.90, p < .05, \eta_p^2 = .06$, respectively]. The main effects of icon familiarity, visual complexity, and function name concreteness were not significant [$F(1,66) = 3.98, MS_e = 1,603,775.11, p = .05; F(1,66) = 1.31, MS_e = 529,471.74, p > .05$; and $F < 1$, respectively]. Interactions with blocks of trials were significant for icon familiarity [$F(6,396) = 6.53, MS_e = 1,144,537.71, p < .01, \eta_p^2 = .09$], semantic distance [$F(6,396) = 51.16, MS_e = 8,967,357.67, p < .01, \eta_p^2 = .44$], and function name familiarity [$F(6,396) = 4.91, MS_e = 860,696.95, p > .05$], but not for visual complexity ($F < 1$) or function name concreteness ($F < 1$).

Table 4
Percentage of Variance Explained and Unique Variance Explained by Each Icon Characteristic and the Total Percentage of Variance Explained by All Characteristics Considered

	Block						
	1	2	3	4	5	6	7
Accuracy							
Semantic Distance							
% Variance explained	53.9**	38.9**	18.3**	19.2**	8.8**	18.2**	8.5**
% Unique variance explained	24.2**	12.3**	2.1	6.0*	0.7	2.6	0.0
Function Familiarity							
% Variance explained	37.5**	40.6**	36.1**	20.2**	19.7**	32.3**	32.0**
% Unique variance explained	7.9**	14.0**	19.9**	7.0*	8.8**	16.8**	23.5**
<i>Total variance explained</i>	61.7	52.8	38.2	26.2	20.4	34.9	32.0
Response Times							
Icon familiarity							
% Variance explained	18.3**	33.6**	35.3**	28.7**	19.2**	11.8**	18.5**
% Unique variance explained	2.7	7.6**	7.9**	7.4**	2.0	0.2	1.9
Semantic Distance							
% Variance explained	54.9**	49.6**	39.0**	27.4**	30.1**	23.6**	21.1**
% Unique variance explained	37.8**	19.5**	9.3**	5.6**	7.8**	5.2*	2.4
Function Familiarity							
% Variance explained	10.3**	25.6**	34.6**	26.2**	29.0**	29.8**	34.0**
% Unique variance explained	1.3	0.6	4.3*	3.3	5.8**	9.9**	11.5**
<i>Total variance explained</i>	58.0	58.5	56.9	43.1	41.3	35.9	39.6

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

The relationship between significant covariates and RTs for significant covariates was examined further, using regression analyses for each block of trials (see Table 4). Semantic distance was the strongest predictor of RTs in early blocks of trials, explaining 55% of the variance initially. The amount of variance predicted by semantic distance, icon familiarity, and function name familiarity was similar in later blocks of trials. The pattern of findings for the unique variance showed that semantic distance was the primary unique predictor of RTs in early blocks of trials. When combined, icon and function familiarity and semantic distance accounted for almost 60% of the variance observed in RTs in early trials and between approximately 35% and 40% of the variance in later trials.

GENERAL DISCUSSION

Naming an object is generally thought to involve three types of stored representations: visual, semantic, and lexical representations. Each form of representation is usually associated with a series of processing steps or stages. Johnson et al.'s (1996) theory is typical of this approach. The processing steps that they delineate are (1) search and perception of the icon (i.e., access to and computation of a visual representation of the object from the visual image), (2) retrieval of a matching representation (i.e., access to stored visual representations), (3) activation of semantic information (i.e., access to conceptual and functional information associated with the object), and (4) access to the function, or name, via referential connections. The findings from our study will now be discussed with reference to this theoretical model.

Our study shows that the stimulus characteristics that are most important in predicting icon identification change as icon–function relationships are learned and present a more dynamic picture than is typically the case for picture naming, where names tend to be overlearned. In line with our expectations, semantic distance was the primary predictor of both accuracy and RT during early learning trials when participants were building connections between icons and their functions, whereas other predictors became important as learning progressed. Johnson et al. (1996) argued that the formation of strong referential links between visual (icon) and verbal (function) was an important predictor of the accuracy and time taken to name pictures. What seems most likely is that semantic distance is an index of the closeness and efficacy of the connections between visual, semantic, and lexical representations.

Function, rather than icon, familiarity predicted unique variance in the participants' identification accuracy throughout the learning trials and, after initial blocks of trials, was the most important predictor of accuracy. For RTs, icon familiarity predicted unique variance in early learning trials, and this is similar to findings from previous research (Isherwood et al., 2007). Isherwood et al. attributed this to participants' ability to use their familiarity with items depicted in the icon in order to ascertain its meaning. In this study, function familiarity became the more important determinant of RTs in later trials. These findings suggest that familiarity with what is in the graphic

is important in early blocks of trials but that naming a familiar object will not always result in an accurate response. Familiarity for both icons (Isherwood et al., 2007) and functions (this study) has a greater long-term effect in determining RTs. The most likely explanation for this pattern of findings is that performance is better for familiar items because long-term memory representations are initially richer and easier to access (Lambon Ralph et al., 1998), and this difference in the quality of the underlying representations remains even after repeated presentations (Jeschaniak & Levelt, 1994). In terms of the processing steps outlined by Johnson et al. (1996), icon familiarity appears to be an index of the ease with which individuals can access stored visual representations (Step 2) and may help to drive initial semantic access (Step 3) if naming the object is sufficient to arrive at the correct function name. Although it is clear that icon and function familiarity overlap considerably in the variance they predict in determining ease of access to semantic representations, function familiarity is also implicated in later processing via its association with function length (Step 4). This is the result of Zipf's (1965) law for function names, where shorter function names tend to be used for more frequently occurring items. This is very much in line with the proposal made by Bates et al. (2003) for frequency effects in picture naming. She and her colleagues suggested that their findings supported the possibility that "frequency effects are not restricted to the word form level. There may be at least two kinds of frequency variance in play: conceptual accessibility . . . and word form frequency" (p. 377).

Visual complexity has been most closely associated with Processing Step 1 (Byrne, 1993; McDougall et al., 2006)—perception and recognition of the icon—although it has also been implicated in retrieval of stored visual object representations (Step 2; Biederman, 1987; Lloyd-Jones & Luckhurst, 2002). In this study, complexity was positively correlated with RTs in later trials (i.e., as complexity increased, RTs increased), but not with accuracy. Subsidiary analyses showed that when complexity was entered as the first item in regression analyses, it explained significant, but modest, amounts of the variance in later blocks of trials (9%–13% of the variance in Blocks 7–13). This pattern of findings is consonant with previous research that has shown that visual complexity effects are associated with the visual search component of the icon identification task (see McDougall et al., 2000) but must be treated with caution, given the flattening of the learning curve and limitations on variation in later blocks of trials.

In earlier research, icon concreteness was often considered the most important characteristic to consider in icon design, since presenting pictorial material that individuals can recognize allows them to access what they know about those objects in order to interpret the icon. However, it was not a major predictor of icon identification, probably because only a limited number of functions can easily be represented pictorially (Rogers & Osborne, 1987; Stammers, 1990; Stammers & Hoffman, 1991). For concepts that are more abstract, it is the closeness of the relationship between icon and function that is important. Because many more concepts can be represented abstractly, icon design needs

to focus more closely on the conceptual mapping between icon and function—that is, on semantic distance. The importance of goodness of fit also seems to be important for picture naming, and where a number of names are possible, this creates uncertainty and slows semantic access and naming RTs (Bates et al., 2003; Johnson et al., 1996).

What is clear from our analysis of the loci of the effects of stimulus variables is that it is difficult to make one-to-one mappings between the effects of a variable and a stage of processing. Similar difficulties have been noted by other authors (Johnson et al., 1996; Lloyd-Jones & Nettlemill, 2007). One explanation for this is that interactivity between forms of representation means that it is difficult to isolate the effects of one variable to a single processing step—not least, because processing outcomes at one level will have an effect on another. Although this study was not explicitly designed to examine the differences between models of naming emphasizing localized stagewise processing and more interactive systems (cf. Levelt et al., 1999, and Nickels & Howard, 1995, with Vitkovitch, Humphreys, & Lloyd-Jones, 1993, and Humphreys & Forde, 2001), the data appear to fit better with an interactive system where one might expect to see the effects of variables fed backward and forward through the processing system. This, however, needs to be explicitly examined experimentally.

It is clear from this research that covariance between variables needs to be taken into account when carrying out and interpreting statistical analyses (see Morrison, 2003, for a useful discussion). Although it has been possible to tease out the predictive role of different variables in this study, it does raise questions about precisely what is being measured when we obtain subjective ratings of stimuli. One possible way around this problem is to develop new measures or metrics that are more objective. For example, objective metrics for visual complexity have been developed (Forsythe et al., 2003; García et al., 1994). Objective measures are also becoming available for semantic distance. Recently, an objective measure of semantic distance between pairs of words was obtained by Maki, McKinley, and Thompson (2004). They combined information from an electronic dictionary, WordNet (Fellbaum, 1998), with co-occurrence and associative norms to produce an objective measure that correlated well with ratings of semantic distance. There are no equivalent dictionaries or electronic databases available for icons, so it may be some time before an objective measure of the semantic distance between icon and function is produced. Ultimately, a better understanding of semantic distance would provide a way of allowing researchers to determine a priori how comprehensible icons are likely to be, at least for individuals learning icon sets.

Another way around the problem of covariance between variables is to combine correlational research, which allows the consideration of several predictors of performance and evaluation of their relative role, with more systematic and orthogonal variation of two or three stimulus characteristics. One variable that it would be possible to investigate using this approach would be age of acquisition. It was not possible to explore its role in this study, because ratings of age of acquisition cannot be sensibly

obtained for icons. However, it would be possible to create new icon sets and vary the time at which icons in the set were introduced (and therefore acquired) over a series of learning trials (see Izura & Ellis, 2002, for the use of such an approach with words). This approach would have the advantage of being able to hold other “nuisance” variables constant, such as frequency and familiarity, which are commonly correlated with age of acquisition.

CONCLUSIONS

It is at least as important to consider predictors of icon identification as picture naming because icons are used so commonly nowadays in public spaces, on vehicles, on machinery, and on computers. Furthermore, it has important practical implications, since it informs icon design. We have drawn attention to the need to consider the nature and characteristics of the function and the goodness of fit between the function and what is depicted in the icon. Highlighting the role of the function label is particularly important, since this, to our knowledge, has received little research attention to date. We have also sought to provide a more theoretically driven view of the processing involved in icon identification, since this has previously been largely ignored by researchers who have primarily been driven to find out “what works” to determine icon usability.

AUTHOR NOTE

This work was supported by an EPSRC/BAE SYSTEMS Industrial Case Studentship (Ref. 98805245). We gratefully acknowledge Neil Carter’s technical and programming support. Correspondence concerning this article should be addressed to S. McDougall, Psychology Department, Bournemouth University, Fern Barrow, Bournemouth BH12 5BB, England (e-mail: smcdougall@bournemouth.ac.uk).

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.
- AREND, U., MUTHIG, K.-P., & WANDMACHER, J. (1987). Evidence for global feature superiority in menu selection by icons. *Behaviour & Information Technology*, *6*, 411-426.
- BARRY, C., MORRISON, C. M., & ELLIS, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. *Quarterly Journal of Experimental Psychology*, *50A*, 560-585.
- 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.
- BIEDERMAN, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, *94*, 115-147.
- BIRD, H., FRANKLIN, S., & HOWARD, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers*, *33*, 73-79.
- BOCKER, M. (1993). A multiple index approach for the evaluation of pictograms. In *Proceedings of the 14th International Symposium on Human Factors in Telecommunication* (pp. 73-84). Heidelberg: Decker’s Verlag, G. Schenck GmbH.
- BOGKA, N., MASTERSON, J., DRUKS, J., FRAGKIUDAKI, M., CHATZ-IPROKOPIOU, E.-S., & ECONOMOU, K. (2003). Object and action picture naming in English and Greek. *European Journal of Cognitive Psychology*, *15*, 371-403.
- BONIN, P., BARRY, C., MÉOT, A., & CHALARD, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory & Language*, *50*, 456-476.

- BONIN, P., CHALARD, M., MÉOT, A., & FAYOL, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, **93**, 89-114.
- BONIN, P., PEEREMAN, R., MALARDIER, N., MÉOT, A., & CHALARD, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, & Computers*, **35**, 158-167.
- BYRNE, M. D. (1993). Using icons to find documents: Simplicity is critical. *Proceedings of INTERCHI '93* (pp. 446-453). New York: ACM Press.
- CARAMAZZA, A., COSTA, A., MIOZZO, M., & BI, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 1430-1450.
- CATLING, J. C., & JOHNSTON, R. A. (2006). Age of acquisition effects on an object-name verification task. *British Journal of Psychology*, **97**, 1-18.
- CUETOS, F., ELLIS, A. W., & ALVAREZ, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, & Computers*, **31**, 650-658.
- CYCOWICZ, Y. M., FRIEDMAN, D., ROTHSTEIN, M., & SNODGRASS, J. G. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology*, **65**, 171-237.
- DELL'ACQUA, R., LOTTO, L., & JOB, R. (2000). Naming times and standardized norms for the Italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, & Computers*, **32**, 588-615.
- 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.
- FAMILANT, M. E., & DETWEILER, M. C. (1993). Iconic reference: Evolving perspectives and an organizing framework. *International Journal of Man-Machine Studies*, **39**, 705-728.
- FELLBAUM, C. (Ed.) (1998). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press. Available at www.cogsci.princeton.edu.
- FEYEREISEN, P., VAN DER BORCHT, F., & SERON, X. (1988). The operativity effect in naming: A re-analysis. *Neuropsychologia*, **26**, 401-415.
- FIEZ, J. A., & TRANEL, D. (1997). Standardized stimuli and procedures for investigating the retrieval of lexical and conceptual knowledge for actions. *Memory & Cognition*, **25**, 543-569.
- FORSYTHE, A., MULHERN, G., & SAWEY, M. (2008). Confounds in pictorial sets: The role of complexity and familiarity in basic-level picture processing. *Behavior Research Methods*, **40**, 116-129.
- FORSYTHE, A., SHEEHY, N., & SAWEY, M. (2003). Measuring icon complexity: An automated analysis. *Behavior Research Methods, Instruments, & Computers*, **35**, 334-342.
- GARCÍA, M., BADRE, A. N., & STASKO, J. T. (1994). Development and validation of icons varying in their abstractness. *Interacting With Computers*, **6**, 191-211.
- GEORGE, D., & MALLERY, P. (2005). *SPSS for Windows step by step: A simple guide and reference, 12.0 update* (5th ed.). Boston: Pearson.
- GILHOOLY, K. J., & LOGIE, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, **12**, 395-427.
- 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* (pp. 457-462). Amsterdam: North-Holland.
- GUASTELLO, S. J., TRAUT, M., & KORIENEK, G. (1989). Verbal versus pictorial representations of objects in a human-computer interface. *International Journal of Man-Machine Studies*, **31**, 99-120.
- HANCOCK, H. E., ROGERS, W. A., SCHROEDER, D., & FISK, A. D. (2004). Safety symbol comprehension: Effects of symbol type, familiarity, and age. *Human Factors*, **46**, 183-195.
- HIRSH, K. W., & FUNNELL, E. (1995). Those old, familiar things: Age of acquisition, familiarity and lexical access in progressive aphasia. *Journal of Neurolinguistics*, **9**, 23-32.
- HUMPHREYS, G. W., & FORDE, E. M. E. (2001). Hierarchies, similarity, and interactivity in object recognition: "Category-specific" neuropsychological deficits. *Behavioral & Brain Sciences*, **24**, 453-476.
- HUMPHREYS, G. W., RIDDOCH, M. J., & QUINLAN, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, **5**, 67-104.
- ISHERWOOD, S. J., MCDUGALL, 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.
- IZURA, C., & ELLIS, A. W. (2002). Age of acquisition effects in word recognition and production in first and second languages. *Psicológica*, **23**, 245-281.
- JESCHENIAK, J. D., & LEVELT, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 824-843.
- JOHNSON, C. J., & CLARK, J. M. (1988). Children's picture naming difficulty and errors: Effects of age of acquisition, uncertainty, and name generality. *Applied Psycholinguistics*, **9**, 351-365.
- JOHNSON, C. J., PAIVIO, A., & CLARK, J. M. (1996). Cognitive components of picture naming. *Psychological Bulletin*, **120**, 113-139.
- KELLER, P., & STEVENS, C. (2004). Meaning from environmental sounds: Types of signal-referent relations and their effect on recognizing auditory icons. *Journal of Experimental Psychology: Applied*, **10**, 3-12.
- LACHMAN, R., & LACHMAN, J. L. (1980). Picture naming: Retrieval and activation of long-term memory. In L. W. Poon, J. L. Fozard, L. S. Cermak, D. Greenberg, & L. W. Thompson (Eds.), *New directions in memory and aging* (pp. 313-343). Hillsdale, NJ: Erlbaum.
- LAMBON RALPH, M. A., GRAHAM, K. S., ELLIS, A. W., & HODGES, J. R. (1998). Naming in semantic dementia—what matters? *Neuropsychologia*, **36**, 775-784.
- LEVELT, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- LEVELT, W. J. M., ROELOFS, A., & MEYER, A. S. (1999). A theory of lexical access in speech production. *Behavioral & Brain Sciences*, **22**, 1-75.
- LLOYD-JONES, T. J., & LUCKHURST, L. (2002). Effects of plane rotation, task, and complexity on recognition of familiar and chimeric objects. *Memory & Cognition*, **30**, 499-510.
- LLOYD-JONES, T. J., & NETTLEMILL, M. (2007). Sources of error in picture naming under time pressure. *Memory & Cognition*, **35**, 816-836.
- MAKI, W. S., MCKINLEY, L. N., & THOMPSON, A. G. (2004). Semantic distance norms computed from an electronic dictionary (WordNet). *Behavior Research Methods, Instruments, & Computers*, **36**, 421-431.
- MARTEIN, R. (1995). Norms for name and concept agreement, familiarity, visual complexity and image agreement on a set of 216 pictures. *Psychologica Belgica*, **35**, 205-225.
- MCDUGALL, 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.
- MCDUGALL, S. J. P., CURRY, M. B., & DE BRUIJN, O. (2001). The effects of visual information on users' mental models: An evaluation of pathfinder analysis as a measure of icon usability. *International Journal of Cognitive Ergonomics*, **5**, 59-84.
- MCDUGALL, S. J. P., DE BRUIJN, O., & CURRY, M. B. (2000). Exploring the effects of icon characteristics on user performance: The role of icon concreteness, complexity, and distinctiveness. *Journal of Experimental Psychology: Applied*, **6**, 291-306.
- MCDUGALL, S. [J. P.], 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.
- MEYER, A. S., SLEIDERINK, A. M., & LEVELT, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, **66**, B25-B33.
- MORRISON, C. M. (2003). Interpret with caution: Multicollinearity in multiple regression of cognitive data. *Perceptual & Motor Skills*, **97**, 80-82.
- MORRISON, C. M., & GIBBONS, Z. C. (2006). Lexical determinants of semantic processing speed. *Visual Cognition*, **13**, 949-967.
- MORRISON, C. M., HIRSH, K. W., & DUGGAN, G. B. (2003). Age of acquisition, ageing, and verb production: Normative and experimental data. *Quarterly Journal of Experimental Psychology*, **56A**, 705-730.

- MOYES, J., & JORDAN, P. W. (1993). Icon design and its effect on guessability, learnability, and experienced user performance. In J. L. Alty, D. Diaper, & S. Guest (Eds.), *People and computers VIII* (pp. 49-60). Cambridge: Cambridge University Press.
- NICKELS, L., & HOWARD, D. (1995). Aphasic naming: What matters? *Neuropsychologia*, **33**, 1281-1303.
- OLDFIELD, R. C., & WINGFIELD, A. (1964). The time it takes to name an object. *Nature*, **202**, 1031-1032.
- OLDFIELD, R. C., & WINGFIELD, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, **17**, 273-281.
- PAIVIO, A., CLARK, J. M., DIGDON, N., & BONS, T. (1989). Referential processing: Reciprocity and correlates of naming and imaging. *Memory & Cognition*, **17**, 163-174.
- PAIVIO, A., YUILLE, J. C., & MADIGAN, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, **76**, 1-25.
- PEIRCE, C. S. (1932). Elements of logic. In C. Hartshorne & P. Weiss (Eds.), *The collected papers of Charles Sanders Peirce* (Vol. 2). Cambridge, MA: Harvard University Press.
- ROGERS, Y., & OBORNE, D. J. (1987). Pictorial communication of abstract verbs in relation to human-computer interaction. *British Journal of Psychology*, **78**, 99-112.
- SANFELIU, M. C., & FERNANDEZ, A. (1996). A set of 254 Snodgrass-Vanderwart pictures standardized for Spanish: Norms for name agreement, image agreement, familiarity, and visual complexity. *Behavior Research Methods, Instruments, & Computers*, **28**, 537-555.
- SANTIAGO, J., MACKEY, D. G., PALMA, A., & RHO, C. (2000). Sequential activation processes in producing words and syllables: Evidence from picture naming. *Language & Cognitive Processes*, **15**, 1-44.
- 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.
- SNODGRASS, J. G., & YUDITSKY, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers*, **28**, 516-536.
- STAMMERS, R. B. (1990). Judged appropriateness of icons as a predictor of identification performance. In E. J. Lovesey (Ed.), *Contemporary ergonomics* (pp. 371-376). London: Taylor & Francis.
- STAMMERS, R. B., GEORGE, D. A., & CAREY, M. S. (1989). An evaluation of abstract and concrete icons for a CAD package. In E. D. Megaw (Ed.), *Contemporary ergonomics 1989* (pp. 416-421). London: Taylor & Francis.
- STAMMERS, R. B., & HOFFMAN, J. (1991). Transfer between icon sets and ratings of icon concreteness and appropriateness. In *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 354-358). Santa Monica, CA: Human Factors Society.
- 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.
- SZÉKELY, A., & BATES, E. (2000). Objective visual complexity as a variable in studies of picture naming. *Center for Research in Language Newsletter*, **12**(2). La Jolla: University of California, San Diego.
- TABACHNICK, B. G., & FIDELL, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: Harper Collins.
- VITKOVITCH, M., HUMPHREYS, G. W., & LLOYD-JONES, T. J. (1993). On naming a giraffe a zebra: Picture naming errors across different object categories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 243-259.
- VITKOVITCH, M., & TYRRELL, L. (1995). Sources of disagreement in object naming. *Quarterly Journal of Experimental Psychology*, **48A**, 822-848.
- VUKELICH, M., & WHITAKER, L. A. (1993). The effects of context on the comprehension of graphic symbols. In *Proceedings of the Human Factors and Ergonomics Society, Forensics Professional* (pp. 511-515). Santa Monica, CA: Human Factors Society.
- WATERS, G. S., SEIDENBERG, M. S., & BRUCK, M. (1984). Children's and adults' use of spelling-sound information in three reading tasks. *Memory & Cognition*, **12**, 293-305.
- WEEKES, B. S., SHU, H., HAO, M., LIU, Y., & TAN, L. H. (2007). Predictors of timed picture naming in Chinese. *Behavior Research Methods*, **39**, 335-342.
- ZEVIN, J. D., & SEIDENBERG, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory & Language*, **47**, 1-29.
- ZIPF, G. K. (1965). *Human behavior and the principle of least effort: An introduction to human ecology*. New York: Hafner.

SUPPLEMENTAL MATERIALS

The icon concreteness and familiarity norms discussed in this article may be downloaded from brm.psychonomic-journals.org/content/supplemental.

(Manuscript received March 4, 2008;
revision accepted for publication November 21, 2008.)