The Use of Multimodal Representation in Icon Interpretation

Siné McDougall[1], Alexandra Forsythe[2], Sarah Isherwood[3], Agnes Petocz[4], Irene Reppa[5], and Catherine Stevens[4]

[1] Bournemouth University, Fern Barrow, Poole, BH12 5BB, UK
[2] Liverpool John Moores University, UK
[3] School of Healthcare, University of Leeds, Leeds, LS2 9JT
[4] University of Western Sydney, South Penrith DC, NSW 1797, Australia
[5] Swansea University, Singleton Park, Swansea, SA2 8PP, UK
smcdougall@bournemouth.ac.uk, a.m.forsythe@ljmu.ac.uk, s.j.isherwood@leeds.ac.uk, {kj.stevens,a.petocz}@uws.ed.au, i.reppa@swan.ac.uk

Abstract. Identifying icon functions differs from naming pictures in that strong semantic links between pictures and their names have been formed over a long period of time whereas the meaning of icons has often to be learned. This paper examines roles of icon characteristics such as complexity, concreteness, familiarity and aesthetic appeal in determining how easily icons can be learned and identified. The role of these characteristics is seen as dynamic, changing as the user learns the icon set. It is argued that the way in which users learn icon meanings is similar to the processes involved in language learning. Icon meanings are learned by drawing on rich multimodal representations which are the result of our world experience. This approach could lead to a better understanding of how multimodal information can be most usefully presented on interfaces.

1 Emerging Themes in Icon Research

Given that communication using icons is now commonplace, it is important to understand the processes involved in icon comprehension and the stimulus cues which individuals utilize to facilitate identification. Earlier research attempted to isolate individual key characteristics which were important in determining such as icon comprehension such as concreteness [1, 2, 3] and visual complexity [4, 5] However, recent work has uncovered a more multifaceted and dynamic picture in which (a) other characteristics such as familiarity, semantic distance and user appeal are important [6, 7, 8] (b) icon characteristics are closely inter-related [e.g. 6, 9], and (c) where the importance of icon characteristics in determining identification changes as the icon set is learned [9]. At the root of effective iconic communication is the way in which the user constructs a relationship between the icon and the function it represents. This has received particular attention when auditory icons are used because, by their very nature, they are more ambiguous and creating an effective icon-referent mapping can be particularly difficult [10, 11, 12, 13].

D. Harris (Ed.): Engin. Psychol. and Cog. Ergonomics, HCII 2009, LNAI 5639, pp. 62–70, 2009.

2 The Effects of Learning on Predictors of Usability

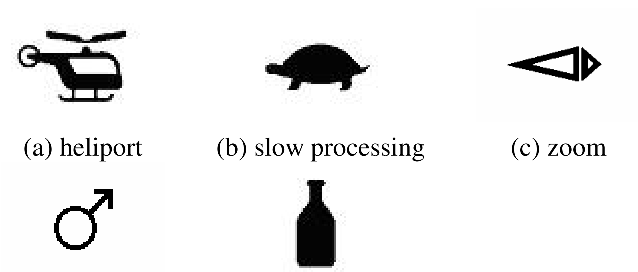
One of the ways in which icons differ from pictures is that strong semantic links between pictures and their names have usually been forged over a long period whereas the intended meaning of an icon often has to be learned. This means that the balance of characteristics which are important in determining usability, and the relationships between those characteristics, changes as users learn icon sets. Changes in the role of each icon characteristic, and their relationship to other characteristics, as learning takes place is now considered.

2.1 Concreteness

Concrete icons which use depictions of real objects and people allow individuals to use their knowledge of the everyday world in order to access meaning (cf Fig. 1a and b with c and d) [2, 3, 14, 15, 16].

Stevens and Petocz [17] have shown that this holds true for auditory, as well as, visual icons. ‘Natural auditory indicators’ such as the sound of a fire engine siren to indicate ‘engine fire’ or the sound of coughing to indicate ‘dangerous levels of carbon monoxide’ allowed participants to use their knowledge of what these sounds normally indicated in order to infer their meaning. Inferring was much more difficult when auditory icons were simple bursts of sound which bore no obvious relationship to the meaning they were indicating. Stevens and Petocz also found that the modality in which icons were presented was important. Individuals responded more accurately and effectively to concrete icons when they were visual than when they were auditory. This appeared to be because auditory icons are generally more ambiguous (e.g. the sound of coughing can indicate many things) and require more interpretation by the user.

Until recently concreteness was sometimes seen as an icon’s most important property, however research now suggests that the effects of concreteness on user performance are less than previously thought [9]. This is because only a limited number of functions can be represented concretely and getting a close fit between pictures and functions is not always easy. For example, naming the bottle in Figure 1e does not allow you to arrive at its intended meaning.



(a) heliport (b) slow processing (c) zoom

(d) male (e) blow moulding

Fig. 1. Examples of different types of icons

Initial differences in performance between concrete and abstract icons are not long lasting and diminish as users gain experience with icons [1, 7, 9]. Other icon properties therefore assume greater importance as learning progresses.

2.2 Semantic Distance

Semantic distance is a measure of the goodness-of-fit between the icon and its meaning. Isherwood and others [9, 18, 19] have shown that semantic distance is more important than concreteness in the initial stages of learning because it is an index of the closeness of icon-function relationships *irrespective* of whether or not the icons are concrete or abstract. This is illustrated in Figure 2. In a study carried out by McDougall, Curry & de Bruijn [20] three sets of icons which were used to indicate functionality on a ‘general purpose vehicle’ designed to overcome obstacles in a road. The abstract and arbitrary icon sets are equally non-pictorial, using primarily shapes and arrows rather than real world items to depict meaning. However, the abstract icons still allowed the user to draw inferences about what might be meant while the arbitrary icons were randomly allocated to functions. Three groups of participants learned the icon sets. Those learning the concrete and abstract sets did *equally* well but those learning the arbitrary set lagged well behind. The latter finding mirrors the findings reported by Stevens & Petocz [16] for their ‘auditory symbolic’ icons in which the meaning of abstract bursts of sounds were only gradually learned by association.

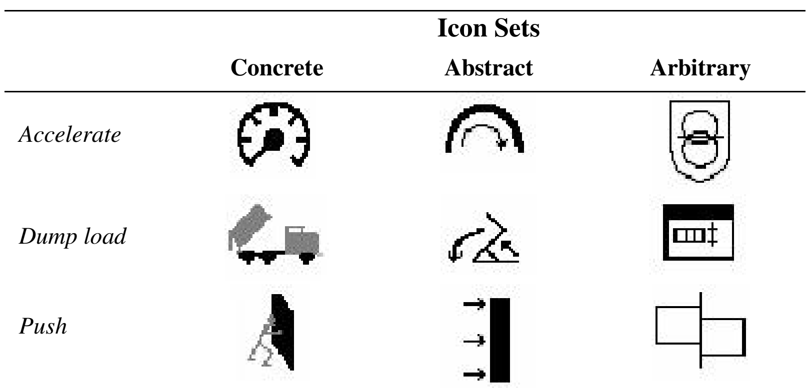


Fig. 2. Examples of icons and functions for concrete, abstract and arbitrary icon sets

2.3 Familiarity

Superficially at least, it seems a truism that familiarity is an important factor in icon interpretation, since familiarity implies something which has already been learned. However, familiarity for icons can take a number of forms; our familiarity with what is depicted in the icon, with the relationship between the icon and its function, and with the function itself. For example, our familiarity with the icon-function relationship allows us to identify the icon representing 'male' in Figure 1e more quickly and effectively compared with the icon representing 'slow processing' (Figure 1b), where we are familiar with what is depicted. Thus, while it can be useful to exploit our familiarity with objects or environmental sounds, if it requires the users to unlearn previously salient meanings such exploitation is likely to have much less utility [17].

Semantic distance is more important than icon familiarity in determining user performance initially as we form relationships between icon and function. Once these have been learned, our familiarity with the icon and the icon-function relationship determine speed and accuracy of responding since it is an index of ease of access to long-term memory [21].

2.4 Complexity

Visual complexity has been most closely associated with the search for, and perception of, icons in displays [4, 7]. The more complex icons are the longer search times on an interface is likely to be. The effects of visual complexity appear to be long lasting with differences in performance between simple and complex icons remaining even after considerable training [7]. Because of its association with visual search, rather than icon identification, its effects on user performance were thought to be relatively independent of other icon characteristics. However, Forsythe's research [22] has shown that this is not the case.

Visual complexity is particularly amenable to the use of metrics and a number of possible measures have been developed. These typically assess a combination of the number of lines and shapes out of which the stimulus is composed. However, Forsythe et al's research [22] has made it clear that these metrics are often biased by our familiarity with the stimulus (more familiar stimuli are perceived as being simpler). For example in the metric developed by Garcia, Badre & Stasko [23], metric calculation is based on icon features including the number of closed and open figures as well as horizontal and vertical lines. It is clear that shapes we are familiar with are more likely to be seen as whole components within the figure, while unfamiliar shapes may be seen as a series of horizontal and vertical lines, so increasingly the complexity value of the stimulus under scrutiny. This 'hidden' relationship between visual complexity and familiarity may account to some extent for the lack of robustness of visual complexity effects with respect to either pictures or icons [24, 25, 26, 27, 28, 29, 30, 31] and suggests that it may have more of a role to play in icon identification than was previously thought.

2.5 Aesthetic Appeal

Research to date for both pictures and icons has shown that there are close relationships between icon characteristics and perceived appeal [32]. Icons that are concrete, familiar and simple have greater appeal in comparison to those that abstract, unfamiliar, and complex [33,34,35,36,37,38]. This evidence might suggest that appeal is simply an emergent property of ease of processing, i.e. when a stimulus is easily and fluently processed this leads to a later positive attribution of this experience as 'liking' for that stimulus [39]. However, the research presented by McDougall, Reppa, Smith & Playfoot [40] suggests that aesthetic appeal has a more active part to play in determining user performance. When users are presented with appealing complex icons performance is better than for unappealing complex icons. Although such effects were not apparent when simple icons were present, this shows that icon appeal can have a direct effect on user performance in a way that the user is unlikely to be conscious of. Such findings therefore go beyond the idea that, if users find an interface appealing they are more likely to spend time learning it, so enhancing performance. Nor is appeal purely stimulus driven. Users underlying cognitive skills also appear to have subtle effects on aesthetic appeal. The direction in which users habitually read influences their preference patterns for icons [40, Study 3].

The trajectory of the effects of aesthetic appeal in determining usability as icons are learned is not yet known but what is clear on the basis of the findings reported in this session is that it is likely to affect usability both in our initial and later encounters with icon sets and that its effects are likely to be closely linked to the effects of other icon characteristics such as concreteness, familiarity and visual complexity.

3 Utilizing Users' Multimodal Knowledge to Enhance Interpretation

While it is important to be able to describe the changing role of icon characteristics as users learning icon sets, contributions to this parallel session [17, 18] have made it clear that the key to enhancing usability is to enable the user to form meaningful connections between the icon and what it is referring to. The closer the icon-function connections formed, irrespective of modality, the more successful the icon is likely to be. What is less clear is the precise form that these connections take, the nature of the representations that underpin them, and the processes involved in learning them. Recent research in psycholinguistics and robotics has led to theoretical developments which may provide a key to understanding the semantic processing and representation underpinning multimodal displays and how they might be optimized.

Grounded cognition theorists argue that semantic information is not stored and processed in an amodal system independent of our sensory systems [41] and that learning takes place by assimilating information multimodally [42]. Grounded cognition theories propose that memory contains many multimodal components from vision, audition, action, affect and language. Retrieving a memory therefore involves reactivating multimodal components together. Such approaches have their roots in early theories of situated action proposed by Gibson and others [43, 44] which have been influential in interface design practice.

Barsalou [41] used the experience of sitting in a chair as an example of how multimodal information capture and subsequent retrieval might work. As we sit down in a chair, the brain captures information across the modalities and integrates them with a multimodal representation in memory. This representation will include, not only how the chair looks, but also how it feels, the creaks it makes as one sits down, the action of sitting, and introspections of comfort and relaxation. Later, when knowledge is needed about a chair (e.g. a chair depicted in a concrete icon to indicate a seating area), our previous multimodal experiences are reactivated. Thus the perception of relevant objects or sounds in icons triggers affordances for action stored in memory [45]. Therefore the extent to which icons are able to activate appropriate memory representations with accompanying appropriate affordances is an important determiner of usability.

Plausible ways in which different types of icon-function relationships might be built are as follows:-

(i) *Direct relationships*: These exploit existing memory representations very straightforwardly. Simply naming the content of the icon provides sufficient information to infer meaning (e.g. the use of helicopter to indicate a heliport in Fig 1a; the use of the sound of a fire engine siren to indicate fire).

(ii) *Indirect relationships*: In these cases the icon has an augmentative function providing an indication of how the meaning might be inferred (e.g slow processing in Fig 1b or the sound of coughing to indicate the build up of noxious gases). Possible referents in memory, though rich in information, are less definite and more diffuse and therefore are more prone to error when the icon-referent relationship is being learned. Abstract icons may benefit particularly if information about likely affordances is indicated in the icon (e.g. the abstract icons in Fig. 2).

(iii) *Distant relationships*: In these cases icon-referent relationships are learned by rote through convention. When initially encountered the icon provides virtually no cues to helpful memory representations, except possibly that this is a type of symbolic information which must be learned and decoded (e.g. the arbitrary icons in Fig. 2; the use of tonal bursts of sound at set frequencies to provide warnings [17]). As the icon-referent relationships are learned they will engender situated memory representations but may never have the same depth or richness of information available to them as direct and indirect relationships.

(iv) *Inappropriate relationships*: When icon-referent relationships cue inappropriate and competing memory representations which cut across their intended meaning, they can be unhelpful. The extent to which this is the case will depend on the richness of the inappropriate memories triggered and will need to unlearn these associations (e.g. 'blow moulding' in Fig. 1e).

Using a grounded cognition approach, we would expect that the meaning of icons is learned in a similar way to the meaning of words and that the process of learning icon-function relationships could be characterized in the same manner. Given this assumption, it is not surprising that the there appear to be strong parallels between children's language acquisition and our ease of learning different types of icon-referent relationships.

Across languages, children's first words consist of object names relating to the items that surround them such as food and toys. This is followed by verbs related to simple actions. Research has shown that the ease with which a word is learned depends upon its concreteness and the extent to which it refers to tangible real world items [46]. Therefore, object names and action verbs are learned first because they are more directly observable and concrete. Later learned words correspond to abstract concepts (e.g. method, justice) that are not directly grounded in experiences with our surroundings. Initial and imageable words directly grounded in physical and perceptual experiences serve as a foundation for the acquisition of abstract words and ideas that become indirectly grounded through their relations to those grounded words (i.e. the use of metaphors with the real world). The capture of multiple stimulus attributes across different modalities to form rich memory representations also helps to explain the probabilistic and complex way in which multiple icon attributes contribute to icon usability and appeal.

One of the difficulties in using multimodal interfaces has been our lack of understanding of precisely how each modality contributes to user experience and vice versa. Given that multimodal learning interfaces are now being built to examine the acquisition of language and meaning using mathematical algorithms [47], this may provide a rich source of information about just how multi-modal information can be balanced to maximally enhance the information provided via interfaces to users.

References

1. Green, A.J.K., Barnard, P.J.: Iconic interfacing: The role of icon distinctiveness and fixed or variable screen locations. In: Diaper, D., et al. (eds.) Human-computer interaction – Interact 1990. Elsevier, North-Holland (1990)
2. Rogers, Y., Oborne, D.J.: Pictorial communication of abstract verbs in related to human-computer interaction. British Journal of Psychology 78, 99–112 (1987)
3. Stotts, D.B.: 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. Human Factors and Ergonomics Society, Santa Monica (1988)
4. Byrne, M.D.: Using icons to find documents: Simplicity is critical. In: Proceedings of INTERCHI 1993, pp. 446–453 (1993)
5. Scott, D.: Visual search in modern human-computer interfaces. Behaviour and Information Technology 12, 174–189 (1993)
6. Forsythe, A., Mulhern, G., Sawey, M.: Confounds in pictorial sets: The role of complexity in basic-level picture processing. Behavior Research Methods 40, 116–129 (2008)
7. McDougall, S.J.P., de Bruijn, O., Curry, M.B.: 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 (2000)
8. Goonetilleke, R.S., Shih, H.M., On, H.K., Fritsch, J.: Effects of training and representational characteristics in icon design. International Journal of Human-Computer Studies 55, 741–760 (2001)
9. Isherwood, S., McDougall, S., Curry, M.: Icon identification in context: The role of icon characteristics and user experience. Human Factors 49, 465–476 (2007)
10. Reppa, I., McDougall, S.: Visual aesthetic appeal speeds processing of complex but not simple icons. In: Proceedings of the 52nd Annual Meeting of the Human Factors and Ergonomics Society 2008, pp. 1155–1159 (2008)
11. McKeown, D., Isherwood, S.: Mapping candidate within-vehicle auditory displays to their referents. Human Factors 49, 417–428 (2007)
12. Petocz, A., Keller, P., Stevens, C.: Auditory warnings, signal-referent relations and natural indicators: re-thinking theory and application. J. Exp. Psychol: Applied 14, 165–178 (2008)
13. Stephan, K.L., Smith, S.E., Martin, R.L., Parker, S.P.A., McAnally, K.: Learning and retention of associations between auditory icons and denotative referents: implications for the design of auditory warnings. Human Factors 48, 288–299 (2006)

14. Arend, U., Muthig, K.-P., Wandmacher, J.: Evidence for global feature superiority in menu selection by icons. Behaviour and Information Technology 6, 411–426 (1987)
15. Stammers, R.B., George, D.A., Carey, M.S.: An evaluation of abstract and concrete icons for a CAD package. In: Megaw, E.D. (ed.) Contemporary ergonomics 1989, pp. 416–421. Taylor & Francis, London (1989)
16. Stammers, R.B., Hoffman, J.: Transfer between icon sets and ratings of icon concreteness and appropriateness. In: Proceedings of the Human Factors Society 35th Annual Meeting, pp. 354–358. Human Factors and Ergonomics Society, Santa Monica (1991)
17. Stevens, K., Petocz, A.: The user knows: Considering the cognitive contribution of the user in the design of auditory warnings. In: Proceedings of HCI International 2009 (2009) (in press)
18. Isherwood, S.J.: Graphics and semantics: The relationship between what is seen and what is meant in icon design. In: Proceedings of HCI International 2009 (2009) (in press)
19. McDougall, S.J.P., Isherwood, S.J.: What's in a name? The role of graphics, functions, and their interrelationships in icon identification. Behaviour Research Methods (in press)
20. McDougall, S.J.P., Curry, M.B., de Bruijn, O.: 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 (2001)
21. Lambon-Ralph, M., Graham, K.S., Ellis, A.W., Hodges, J.R.: Naming in semantic dementia: What matters? Neuropsychologia 36, 775–784 (1998)
22. Forsythe, A.M.: Visual complexity: Is that all there is? In: Proceedings of HCI International 2009 (2009) (in press)
23. Garcia, M., Badre, A.N., Stasko, T.: Development and validation of icons varying in their abstractness. Interacting with Computers 6, 191–211 (1994)
24. Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U.H., Segui, J.: Predictors of picture naming speed. Behavior Research Methods, Instruments & Computers 36, 140–155 (2004)
25. Humphreys, G.W., Riddoch, M.J., Quinlan, P.T.: Cascade processes in picture identification. Cognitive Neuropsychology 5, 67–103 (1988)
26. Ellis, A.W., Morrison, C.M.: Real age-of-acquisition effects in lexical retrieval. Journal of Experimental Psychology: Learning, Memory & Cognition 24, 515–523 (1998)
27. Barry, C., Morrison, C.M., Ellis, A.W.: Naming Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. Quarterly Journal of Experimental Psychology 50A, 560–585 (1997)
28. Bonin, P., Peereman, R., Malardier, N., Méot, A., Chalard, M.: 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 (2003)
29. Cuetos, F., Ellis, A.W., Alvarez, B.: Naming times for the Snodgrass and Vanderwart pictures in Spanish. Behavior Research Methods, Instruments and Computers 31, 650–658 (1999)
30. Paivio, A., Clark, J.M., Digdon, N., Bons, T.: Referential processing: Reciprocity and correlates of naming and imaging. Memory & Cognition 17, 163–174 (1989)
31. Snodgrass, J.G., Yuditsky, T.: Naming times for the Snodgrass and Vanderwart pictures. Behavior Research Methods, Instruments, & Computers 28, 516–536 (1996)
32. McDougall, S., Reppa, I.: Why do I like it? The relationship between icon characteristics, user performance and aesthetic appeal. In: Proceedings of the 52nd Annual Meeting of the Human Factors and Ergonomics Society 2008, pp. 1257–1261 (2008)

33. Berlyne, D.E.: Studies in the new experimental aesthetics. Hemisphere, Washington (1974)
34. Jacobsen, T., Hofel, L.: Aesthetic judgments of novel graphic patterns: Analyses of individual judgments. Perceptual & Motor Skills 95, 755–766 (2002)
35. Bauerly, M., Liu, Y.: Computational modeling and experimental investigation of the effects of compositional elements on interface and design aesthetics. In: Proceedings of the 8th Annual International Conference on Industrial Engineering – Theory, Applications and Practice (2003)
36. Vartanian, O., Goel, V.: Neuroanatomical correlates of aesthetic preferences for paintings. Neuroreport 15, 893–897 (2004)
37. Kawabata, H., Zeki, S.: Neural correlates of beauty. Journal of Neurophysiology 91, 1699–1705 (2004)
38. Bornstein, R.F.: Exposure and affect: Overview and meta-analysis of research, 1968-1987. Psychological Bulletin 106, 265–289 (1989)
39. Reber, R., Schwarz, N., Winkielman, P.: Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? Personality and Social Psychology 8, 364–382 (2004)
40. McDougall, S., Reppa, I., Smith, G., Playfoot, D.: Beyond emoticons: Combining affect and cognition in icon design. In: Proceedings of HCI International 2009 (2009) (in press)
41. Barsalou, L.W.: Grounded cognition. Annual Review of. Psychology 59, 617–645 (2008)
42. Clark, A.: Being there: Putting brain, body, and world together again. MIT Press, Cambridge (1997)
43. Lakoff, G., Johnson, M.: Metaphors we live by. University Chicago Press, Chicago (1980)
44. Gibson, J.J.: The ecological approach to visual perception. Houghton Mifflin, New York (1979)
45. Glenberg, A.M., Schroeder, J.L., Robertson, D.A.: Averting the gaze disengages the environment and facilitates remembering. Mem. Cogn. 26, 651–658 (1998)
46. Gillette, J., Glietman, H., Glietman, L., Lederer, A.: Human simulations of vocabulary learning. Cognition 73, 135–176 (1999)
47. Yu, C., Ballard, D.H.: A multimodal learning interface for grounding spoken language in sensory perceptions. ACM Transactions on Applied Perceptions 1, 57–80 (2004)