People & Computers, 1986, Proc. of HCI'86, Eds M. Harrison & A. Monk.

EVALUATING THE MEANINGFULNESS OF ICON SETS TO REPRESENT COMMAND OPERATIONS

Yvonne Rogers
Department of Psychology, University College Swansea,
Singleton Park, Swansea, West Glamorgan, SA2 8PP, Wales.

Iconic interfacing is becoming increasingly popular as a medium to present information about computer systems and their command operations. This paper considers the extent to which various icons that differ in the form of correspondence between referent and icon symbol can effectively represent a large number of abstract command operations typically used in a word processing environment.

Six icon sets, depicting either abstract symbols, concrete objects operated on, concrete analogies associated with the action or combinations of these were constructed to represent 20 commands covering a range of word processing operation areas. Using a questionnaire, 60 subjects (10 for each set) were required to match the icons to the commands they thought they referred to. Significant differences were found between the icon sets. Specifically, the icon sets with the most direct mapping (i.e. those depicting concrete objects operated on) were found to have the highest number of correct matches, with over 85% of the icons being correctly identified. An interaction between icon set and type of command was also found indicating that some commands can be represented in a range of pictorial forms while for other commands the type of pictorial form is critical. The results from this experiment are discussed in relation to the demands made on the cognitive resources for the comprehension of visual symbols.

1 INTRODUCTION

One of the major challenges in Human-Computer-Interaction is to design an interface which is simple enough to learn, yet complex enough

to allow a diversity of operations to be performed. Recently, there has been a growing trend to develop interfaces which incorporate icons and various other visual symbols: the idea being that visual symbols can help reduce the complexity of the system to the user and make it easier to understand and use (Lodding, 1983).

Underlying this idea is the popular notion that since we live in a strongly visual and spatially organised environment it seems more compatible to learn interfaces which also use visual and spatial information representations. For example, icon-driven office automation systems such as Xerox Corporation's 'Star' (Smith, Irby, Kimball and Verplank, 1982) and Apple Corporation's 'Lisa' (Williams, 1983) both claim that an icon system is able to provide the user with a visual image which acts as a physical metaphor of the real world. By creating electronic analogues, in the form of concrete objects, to the physical objects of an office (e.g. icons representing papers, folders, filecabinets and wastepaper bins) they propose that this should make the electronic 'world' seem less alien and require less training.

Similar benefits have also been proposed for ICADS which is a computerised icon-driven architectural modelling and drafting system (Cornell, Sambura and Gero, 1984) and for spatial data management systems where data are presented to the user in a concrete iconic form (Herot, 1980). In addition, these applications emphasise the idea that iconic interfaces are actually enjoyable to use and it is this fun element which can provide the initial incentive to learn how to use them.

Besides the attractive appeal to novices, graphic images in general are often regarded as a potentially universal means of communication and are assumed, therefore, to be able to overcome some of the problems associated with verbal language. In particular, certain types of information can sometimes be conveyed more directly and with more immediacy than a verbal equivalent. Hemenway (1982), for example, suggests that the advantage of icon systems over verbal representations are similar to the benefits of other graphic systems such as maps. Compared with a set of words she suggests graphic symbols can be visually more distinctive from one another and can represent variation within a set of commands more effectively. Rohr and Keppel (1984) also propose that using icons enable certain types of complex information to be presented in a more spatially condensed and holistic form.

One of the main advantages of iconic interfacing, therefore, would appear to lie in its potential to represent a large variety of commands and functions that are easy to identify. An obvious implication, therefore, is that many of the initial difficulties often experienced in learning and subsequently remembering command names (see for example Long, Hammond, Barnard and Morton, 1983) could be overcome.

A problem with substituting icons for command names, however, is that the demands placed on the cognitive resources required to learn them may prove actually to be greater than those required for the equivalent verbal set. Although a set of icons has the potential to impart more knowledge about the system by being more spatially condensed than an equivalent set of verbal commands, it may be at the expense of making extra demands on the user's memory resources and abstracting ability. The purpose of this study, therefore, was to investigate the initial comprehensiveness of icons designed to represent a variety of set of command operations, with the aim of determining whether a large number of functions could be effectively mapped into a graphic form that was easily discriminable.

Regarding the extent to which abstract concepts are meaningful when represented in a pictorial form, it appears from the psychological literature that little is known about how visual symbols representing abstract concepts are processed and, moreover, how well they can be understood. Since it becomes more difficult to represent abstract concepts the more abstract they become, there is a strong preference for depictions which attempt to concretize the concept in some way. (for example, Jones, 1983). Howell and Fuchs (1968), however, who earlier had considered abstract military intelligence concepts found that classes of abstract concepts defined as actions might best be represented visually either by using a combination of concrete objects with abstract symbols or through an arrangement of pictorial symbols from which the interaction sequence can be deduced. For example, the most effective form of visual representation were symbols which consisted not only of armoured vehicles and factory buildings but which were also accompanied by conventional abstract symbols such as ticks, crosses and various forms of lines. These results suggest, therefore, that computer command operations which have underlying action primitives would, likewise, be represented more meaningfully in this manner.

Hemenway (1982) has also suggested that the 'directness' of the link between what is depicted in the icon and the command is likely to affect both the initial comprehension and subsequent retention. She proposes

that the mappings which have the most direct link i.e. icons which depict objects or operations directly involved in the commands should be easier to initially comprehend than those that represent the command by analogy. There appears as yet, however, no empirical evidence to support this theory.

The other aim of this study, therefore, was to evaluate the differences in meaningfulness for various types of mappings between command and icon, using either concrete or abstract image elements or a combination of both. It was hypothesised that icon sets depicting concrete objects operated on in conjunction with abstract symbols denoting the state or direction of the action would be the most effective way of portraying command operations.

As well as investigating the type of pictorial representation, the study was also designed to evaluate the effects of the type of function underlying the meaning of the command. Previous research has shown that properties underlying the meaning of command verbs have an effect on user performance in memory tasks involving computer systems (Rogers and Oborne, 1985). Furthermore, Rohr (1984) has shown that commands concerned with processing functions (e.g. "print") were rated as being more meaningful when pictorial representations consisted of a prototypical object performing this process. Commands requiring manipulative control actions (e.g. "insert"), however, were best represented by a prototypical action sequence comprising of abstract visual symbols. Thus, it was further hypothesised that the function and outcome of the operation would have an effect on the efficacy with which the command could be pictorially represented.

The purpose of this study, therefore, was first to investigate the initial comprehensiveness of icons designed to represent a large number of command operations typically used in a word processing environment. Secondly, to determine whether there was any difference in meaningfulness between concrete and abstract representations or their combinations to depict various types of command operations. Thirdly, to invesigate the extent to which the directness of the link between the pictorial representation and the command operation is important for the comprehension of the icon.

2 METHOD

The method used to measure the comprehensiveness of the icons was a 'matching test'. The basis of this type of test is that a respondent is asked to select for a given referent one from a number of symbols. An index of comprehensiveness is then derived from the number of times a symbol is correctly selected from the group of symbols presented to the respondents. Subsequently this score is related to the set of symbols presented simultaneously. In addition, incorrect selections of symbols i.e. the confusions between referents and symbols can be used as a measure. This type of method has previously been used by Krauss as early as 1930 and by Scheerer and Lyons in 1957 (cited in Werner and Kaplan, 1967, p. 341) to investigate the understanding of graphically expressed abstract concepts. More recent research which has used this type of procedure include Easterby and Zwaga (1976) and Zwaga and Boersema (1983) on public information symbols.

2.1 Subjects

The subjects were 60 (10 for each icon set) unpaid students at University College Swansea. There were 33 females and 27 males, whose ages ranged between 18 and 26.

2.2 Materials

2.2.1 Stimuli

A questionnaire was designed which consisted of 20 word processing commands and a set of 20 icons. Four types of commands representing the range of operation areas found in a typical word processing environment were used. The categories of commands and their functions together with examples of each are provided in Table 1.

Table 1 - Categorisation and examples of command operations

CATEGORY	FUNCTION	EXAMPLE
text movers	moving cursor around the text	go to top of text
text manipulators	change specific parts of text	insert new line
	change blocks of text	delete block of text
file processors	carry out operations before or after typing text	print a file of text
state controllers	carry out operations in operating mode	show a directory

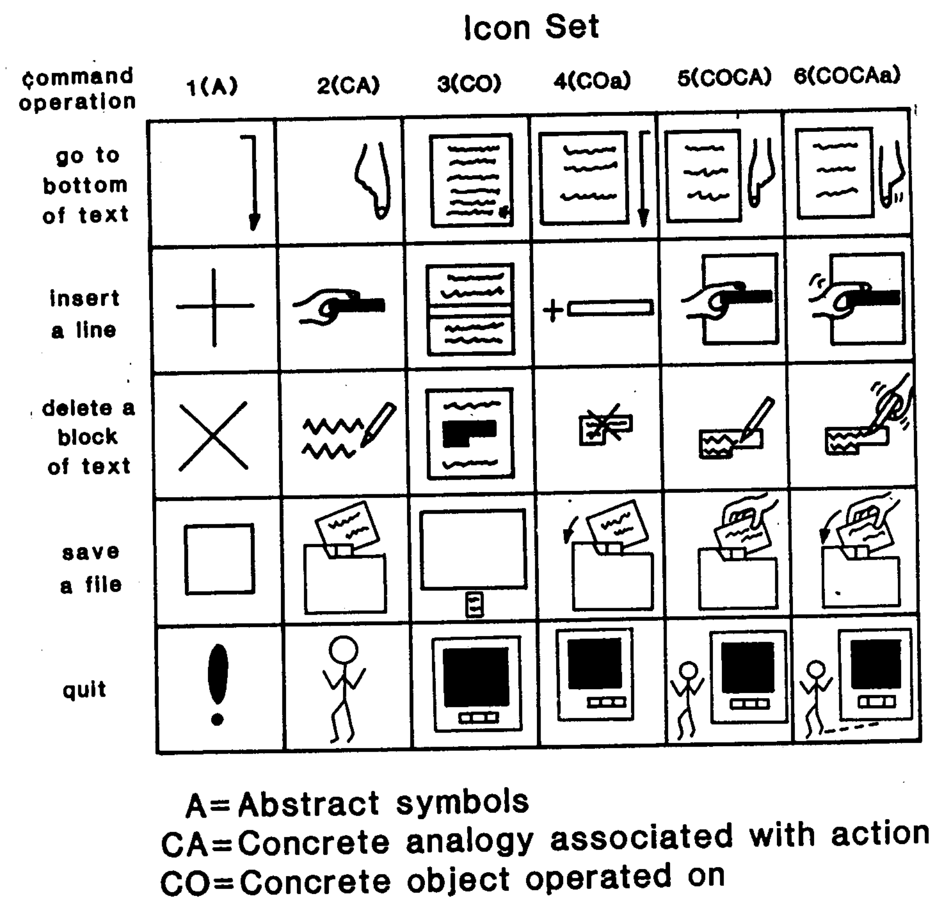
2.2.2 Icon Sets

Six different sets of icons were constructed to represent the commands. They were categorised in terms of being abstract, concrete or a combination of both and were ordered along a simple-complex scale (see Table 2). The images used in the icons were drawn in black and white and were consistent across sets. The elements used in the images consisted of either abstract symbols e.g. arrows, crosses, shapes and various lines; concrete objects operated on e.g. files, paper and sections of text; concrete analogies associated with the action e.g. a pointing finger, a marker flag and a hand holding a pen; or various combinations of these. They were drawn by a professional artist who was provided with guidelines on what was considered to be prototypical images for the objects and symbols. These were based on findings derived from an earlier unpublished study which investigated the stereotypy of pictorial representations produced for a variety of command verbs. Examples of the icons constructed for the six sets are shown in Figure 1.

Table 2 - Description and categorisation of icon sets

ICON SET	DESCRIPTION	CATEGORY	EXAMPLES OF ELEMENTS (for the command 'go to bottom of text')
1	abstract symbol	simple-abstract (A)	arrow
2	concrete analogy	simple-concrete (CA)	pointing finger
3	concrete object operated on	simple-concrete (CO)	sheet of paper with mark on
4	concrete object plus abstract symbols denoting analogy	complex-concrete, abstract (COa)	arrow pointing to bottom of sheet of paper
5	concrete object plus concrete analogy	complex-concrete, concrete (COCA)	pointing finger to bottom of sheet of paper
6	concrete object plus concrete analogy plus abstract symbols	complex-concrete, concrete abstract (COCAa)	finger with lines denoting movement pointing to bottom of sheet of paper

Figure 1 - Examples of icons used in the Word Processing Matching Task



2.2.3 Questionnaire

Booklets consisting of one icon set and, separately, the set of command operations were constructed. The icons measuring 25 x 25 mm, were randomly presented in a 5 x 4 matrix at the front of each booklet. Above each icon a letter was placed so that it could be identified.

The second section of the booklets consisted of the command operations. These were ordered into five sections which corresponded with the selected categories but were labelled differently so as to make them easier to understand. The following headings were used:-

1] moving around the text
2] making changes to the text
3] making changes to blocks of text
4] operating commands
5] general commands

Under each heading a description of the type of commands included in each section was provided. To record the subject's answer the sentence 'I think icon . . . represents this' was presented by each command.

At the beginning of the questionnaire a scenario was provided with the intention of familiarising the subjects with an appropriate context. It began by asking the subjects to imagine that they have just finished writing a long essay and when reading through it find a number of mistakes in it. After making various changes they are then told that the essay looks a complete mess and so they have to resign themselves to rewriting it. The scenario continues by telling them that a friend of their's invites them to look at a new computer called a word processor which s/he has just bought. The friend explains what the word processor can do. It is then suggested to the subjects that this sounds just what they need for helping them write their essay. The friend agrees to let them have a go and sets it up. S/he gives a brief description of how it works and then leaves it to the subjects. The scenario continues by explaining that on the computer display in front of them are a set of symbols called icons which refer to the various word processor's editing functions and commands. Pointing to one will result in the commands being executed. The subjects, however, have not been told what they represent and so have to infer what they mean.

2.3 Procedure

Subjects were instructed to match the icons to the command operations which they thought appropriate. They were asked to use only one icon for each command and to cross out each icon after they had used it. No time limit was set but, on average, subjects tended to take 15-20 minutes to complete the task.

3 RESULTS

The data were initially analysed for the number of correct matches made for each icon set. The two icon sets depicting objects operated on, with and without the inclusion of abstract symbols, had the highest number of correct matches with over 85% of the icons being correctly identified. The icon set which was found to be the most difficult to match to the commands, on the other hand, was the abstract symbol set where just over half of the icons were correctly identified.

A one-way analysis of variance (ANOVA) performed on the data revealed a main effect for icon set ($F_{5,54} = 10.58$, $p < 0.001$). To determine the significance between icon sets, a post hoc comparison between the treatment means was then performed. Using the Tukey test it was found that there were a number of significant differences between the treatment means ($\overline{d}_T = 4.26$, $p < 0.5$). A summary of these are provided in Table 3. As can be seen, the largest significant differences are found when comparing icon set 2 (concrete analogy) with icon set's 3 (concrete object), 4 (concrete object plus abstract symbol) and 5 (concrete object plus concrete analogy), respectively. The differences obtained show that the means for the icon sets consisting of concrete objects operated on, alone or in combination with other elements, are significantly greater than that obtained for the icon set depicting concrete objects through analogy. Thus it can be inferred that the depiction of concrete objects operated on is by far the most direct and meaningful way of representing abstract referents.

Table 3 - Post hoc comparison between icon set means for correct number of matches

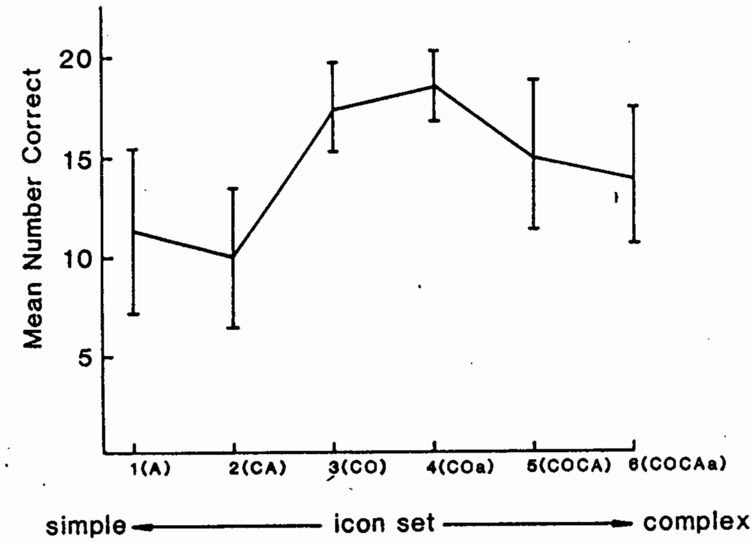
	ICON SETS					
	S2	S1	S6	S5	S3	S4
MEANS	9.9	11.3	13.9	14.9	17.3	18.5
S2		1.4	4.0	5.0**	7.4**	8.6**
S1			2.6	3.6	6.0**	7.2**
S6				1.0	3.4	4.6*
S5					2.4	3.6
S3						1.2

*p < 0.05   ** p < 0.01

Interestingly, the results also show that the icon set depicting objects operated on in combination with abstract symbols was superior to the icon set depicting just concrete objects. This suggests, therefore, that the meaning of the abstract command is more effectively portrayed when some form of additional abstract representation is used e.g. an arrow, in combination with a concrete object to indicate the state of the command. The results also indicate, however, that the addition of a concrete analogy to this combination is actually detrimental in that it becomes more difficult to comprehend icons which have the combination of all three elements. This seems to imply, therefore, that clarity and meaningfulness begin to diminish in highly complex representations. The same can also be said, however, about the simple and abstract representations at the other end of the scale since they appear to be equally inadequate at effectively representing these type of abstract concepts.

Taken together, these results suggest that the data can be roughly fitted to an inverted U shape distribution, whereby the most simple and most complex representations are seen to be the least effective and the representations in the middle the most effective. This trend can be seen in Figure 2 which shows the means and standard deviations for icon sets as a function of number of correct matches.

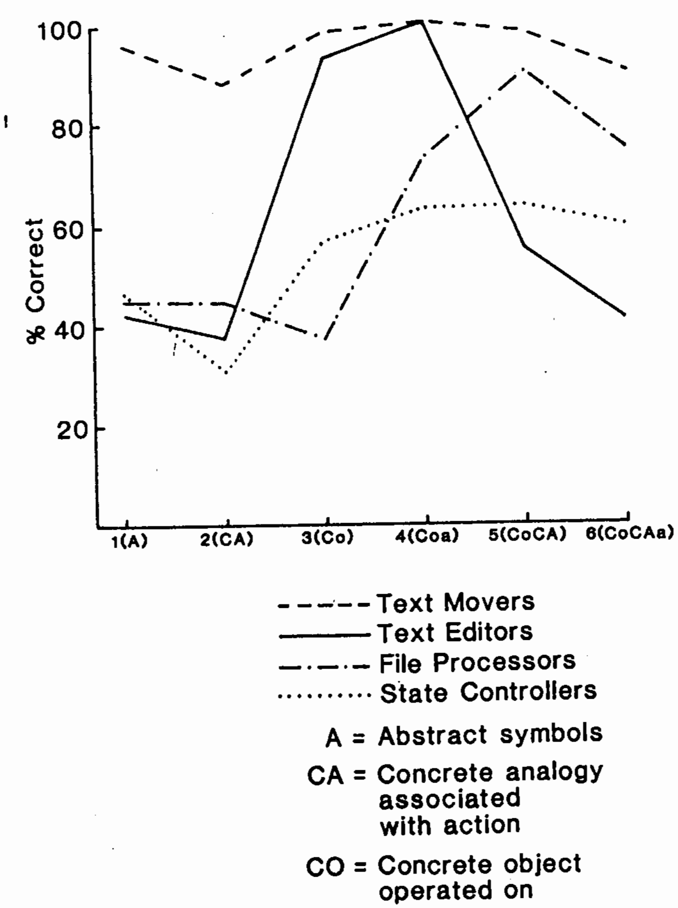
Figure 2 - Mean and S.D. of number of correct matches made for each icon set (n = 10)



A = Abstract symbols
CA = Concrete analogy associated with action
CO = Concrete object operated on

A between subjects analysis of the individual commands also revealed a main effect for icon set ($F_{5,114} = 8.36$, $p < 0.001$). Moreover, an interaction between icon set and command type ($F_{15,162} = 3.66$, $p < 0.001$) was found using a 2-way ANOVA. Figure 3, which shows the mean number of correct matches as a function of command type, illustrates the nature of this interaction. As can be seen from the profile the "text mover" commands were found to be the easiest to understand in pictorial form, independent of type of icon. Conversely, the group of commands categorised as 'state controllers' were found to be relatively less easily understood across all icon forms. On the other hand, the other two categories showed rather different patterns of response in that both produced a differentiated but distinctive effect across icon sets. The 'file processors' showed a 'step-like' effect whereby the commands were difficult to understand when represented by the simple icon set (1-3) but greatly improved when represented by the more complex icon sets (4-6). The 'text editors', on the other hand, showed a similar inverted U shape to that illustrated for the overall result.

Figure 3 - Interaction profile of mean percentage correct matches for command types as a function of icon set



These findings indicate, therefore, that some command operations can be represented by a range of pictorial forms that can equally portray their intended meaning, whilst for other command operations the type of pictorial form is critical. In addition, command operations which appear to be the most abstract i.e. the state controllers, appear to be very difficult to understand in any pictorial form.

The problems involved in understanding the iconic representations of individual command operations were further analysed in terms of the number and type of confusions between icon and referent. A criterion of 3 or more incorrect choices for a referent was decided as an indication of a relevant confusion. Based on this, it was found that there was a total of 34 relevant confusions, which were distributed among the icon sets as follows:-

Set	No. confusions
1	4
2	12
3	3
4	1
5	7
6	7

The confusions were then qualitatively analysed using Zwaga and Boersema's (1983) classification of errors as a guideline.

One type of error which commonly occurred was a symmetrical interpretation confusion whereby elements from one image were misinterpreted for elements of another and vice versa. In particular, several of these confusion errors were found for the icons depicting the insert and delete command operations in the concrete analogy sets. For instance, several of the subjects confused the image elements representing the 'delete a line' and 'delete a space' commands with the image elements representing the 'insert a line' and 'insert a space'. Another symmetrical interpretation confusion which occurred was for the command oeprations 'save a file' and 'retrieve a file'. The reason for this seemed to be caused by a difficulty in differentiating between the two states when represented by similar elements. It was only in the icon sets which included additional elements to indicate the direction of the movement in the image that the correct interpretation was made easier.

A number of assymetrical confusions were also found in the data. For example, there were occasions when the icons for 'destroy a file' were selected for the command 'show a directory' but not vice versa. Alternatively, several subjects selected the icon designed for 'delete a block' or 'to quit' for the command 'to destroy a file'. Similar to the symmetrical confusions, these type of errors may be based on the misinterpretation of image elements. Alternatively, however, it may be that they are actually artefacts of the experimental design. The reason for this is because when a subject has selected one or two incorrect icons, they are then left in the situation of having to make the following matches from the remaining pool of icons. Consequently, they have to choose an alternative icon and may, therefore, select one which has similar features to the correct icon but which represents a different command in that category. The effect of this type of strategy could result in a 'chaining sequence' where a series of confusions are made because the icon set is unable to adequately differentiate between the different operations.

This type of artefact, however, did not appear to occur for the icon sets depicting concrete objects operated on. The very high rate of correct matches for these sets, suggests instead that the icons depicted in these sets were able to effectively discriminate among the different commands.

4. DISCUSSION

The main result from this study was to show that when asked to match a set of icons to word processing commands, a large variety of command operations were able to be comprehended successfully in an iconic form. The extent to which the iconic forms were meaningful initially, hwoever, was found to be dependent both on the directness of the mapping between the pictorial representation and referent and the type of image elements portrayed. In particular, the findings showed that the most effective type of mapping was of pictorial representations depicting concrete objects directly operated on in the referent (CO). Conversely, the icons which were shown to be the least effective were those which depicted the command operation by analogy (CA).

These findings suggest, therefore, that the task of discovering the link between icon and referent is a relatively simple process when the icon directly depicts the operation involved in the command. However, this task becomes increasingly more difficult as the level of correspondence between the two becomes less direct. In terms of the cognitive processing involved, therefore, it would appear that different demands are being made on the resources.

A theoretical explanation for this is that the mental representations resulting from the information accessed from the icons in the CO category are sufficient enough for them to be mapped directly onto their intended

cognitive resources are minimal. The icons in the CA category, however, are unable to provide adequate information for the process of direct mapping. Instead access to additional cognitive resources which can connect the features of the command with the features of the familiar entities used in the analogy are necessary. The knowledge required to draw these parallels, however, may be restricted or not available. Consequently much greater demands will be made on the inferential processes in order that a putative interpretation of the icon in the context of the other icons be made.

In such circumstances there is likely to be an increase in performance costs. For example, one such effect which was found in the study was in the distribution of confusion errors among the various icon categories. Compared to the relatively few errors shown in the CO categories, there was a significantly greater number found in the CA categories. Another example of this effect found in the study was the adoption of a 'chaining sequence' strategy as indicated by some of the subject's results in the CA conditions.

These types of errors seem to suggest that there is some kind of mismatch in the CA conditions between the mental representation resulting from the icon and the cognitive processing required to infer correctly its intended meaning. One reason for this may be due to the inherent ambiguity of analogies. For while such analogies are directly able to convey what they literally represent, they may not be able to subsequently convey their symbolic message. For example, the icon depicting a person with their arms in the air is intended to portray literally a typical gesture commonly associated with giving up. The familiar entities in the image are then supposed to be projected into the word processing domain such that they suggest the specific message of 'to quite the system'. If the properties in the analogy, however, are not particularly salient then an ambiguous situation may arise where alternative messages may be equally interpreted. Thus instead of comprehending the message 'to quit' subjects might readily associate this analogy with other messages such as 'I don't know' or 'help'.

One of the dangers of these type of icons, therefore, is their propensity to be naturally imprecise. If this is the case, then in order to understand the meaning of them it may be necessary to refer to a manual. As argued by Jervell and Olsen (1985) the necessity of such a changed into a symbol and hence the advantage claimed for it being easier to comprehend, initially is no longer valid. It may still be the case, however, that the cognitive resources required to learn such icons and subsequently to remember the commands they refer to are less than those required to learn a verbal equivalent.

To determine whether the use of such icons are actually beneficial at the interface, therefore, it would be necessary to carry out further research comparing the various performance costs that occur in the comprehension, learning and remembering of different visual and verbal representations. This research would also be valuable in providing further insight in to the extent of the efficacy of the icons in the CO conditions.

From these initial results it would, however, appear that this type of representation is able to avoid many of the ambiguity problems associated with the CA conditions. Furthermore, the finding that initial comprehension was significantly increased when the use of abstract symbols were added to the CO condition suggests that it may be possible to further reduce the number of alternative meanings a pictorial representation conveys.

Finally, the other important factor which was found to effect the communicativeness of the icons in the task was the actual function underlying the commands. A general trend which appeared to emerge was that command operations considered to have a high functional content and/or a definite visible outcome e.g. the text movers, were much easier to comprehend in any pictorial form compared to those judged to be low in functionality and to have a less visible outcome e.g. the command 'to save a file'. Interestingly, an interaction also occurred whereby certain command operations were only able to be understood effectively in the CO conditions. What this seems to imply, therefore, is that the more abstract the function is the more difficult it is to comprehend the command operation in pictorial form. Alternatively, the less abstract the command is, the easier it is to represent in a meaningful pictorial form.

In summary, the findings from this initial 'paper and pencil' study suggest that the success of an icon to convey its intended meaning will depend on the combination of a number of factors. In particular, these include the type of pictorial repreentation used, the underlying functionality of the command operation and above all how directly these

5. REFERENCES

Cornell, D., Sambura, A. & Gero, I. (1984) Icon-driven interfaces for drafting systems. In 'Computer Graphics '84', Proceedings of the International Conference, pp. 131-140. London, England.

Easterby, R.S. & Zwaga, H.J.G. (1976) Evaluation of public information symbols: ISO tests, 1975 series. AP Report No. 60. Applied Psychology Department, University of Aston, Birmingham.

Hemenway, K. (1982) Psychological issues in the use of icons in command menus. In 'Proceedings on Human Factors in Computer Systems', pp. 20-25. Gaithersburg, Maryland, March 15-17.

Herot, C.F. (1980) A spatial graphical man-machine interface. In 'Information Processing 80', ed. S.H. Lavington, pp. 1039-1044. North Holland Publishing Company.

Howell, W.C. & Fuchs, A.F. (1968) Population stereotypy in code design. Organisational Behavior and Human Performance, 3, 310-339.

Jervell, H.R. & Olsen, K.A. (1985) Icons in man-machine communications. Behaviour and Information Technology, 4, (3), 249-254.

Jones, S. (1983) Stereotypy in pictograms of abstract concepts. Ergonomics, 26, 605-611.

Long, J., Hammond, N., Barnard, P. & Morton, J. (1983) Introducing the interactive computer at work: The user's views. Behaviour and Information Technology, 2, 39-106.

Rohr, G. (1984) Understanding visual symbols. In 'Proceedings on IEEE Computer Workshop on Visual Languages', pp. 184-191. Hiroshima, Japan 6-8 December.

Rohr, G. & Keppel, E. (1984) Iconic interfaces: Where to use and how to construct. In 'Human Factors in Organisational Design and Management, pp. 269-275. North Holland: Elsevier Science Publishers.

Rogers, Y. & Oborne, D.J. (1985) Some psychological attributes of potential computer commands. Behaviour and Information Technology, 4, (4), 349-365.

Smith, D.C., Irby, C., Kimball, R. & Verplank, B. (1982) Designing the Sta user interface. BYTE, 7, (4), 242-282.

Werner, H. & Kaplan, B. (1967) Symbol Formation. New York: John Wiley.

Williams, G. (1983) The LISA computer system. BYTE, 8, (2), 33-50.

Zwaga, H.J. & Boersema, T. (1983) Evaluation of a set of graphic symbols.