User–Computer Interaction: Cognitive Properties of Icons for Multidimensional Data Analysis

Antonio Messina[1], **Laura Moltedo**[2], **Silvana Contento**[3], **Roberto Nicoletti**[4]

[1] *Dipartimento di Fisica "A. Righi", Università di Bologna, Bologna, Italy*
[2] *Istituto per le Applicazioni del Calcolo "M. Picone" - C.N.R., Roma, Italy*
[3] *Dipartimento di Psicologia, Università di Bologna, Bologna, Italy*
[4] *Dipartimento di Scienze Biomediche, Università di Modena, Modena, Italy*

Abstract

In the graphical environment, icons (i.e. symbols or pictorial representations analogous to physical objects, actions or functions) are generally more informative and easier to manipulate than verbal labels. Because icons should resemble the functions they represent, pre–testing should be done to determine whether the user understands what the icons are supposed to symbolize. In the present work we aimed at establishing whether the criteria adopted to shape the icons of a graphic tool for the analysis of multidimensional field (MUDI3) were shared by the user. Method, procedure and results are presented and discussed.

1 Introduction

The result of the dramatic change occurred during the recent years in the field of computing has been the sudden growth in the interest in distributed computing environments which consist of many autonomous computers of varying computing power connected by networks.

Although enabling researchers to envisage and realize new approaches to computer modeling and analysis of physical systems and data, such an environment requires the development of efficient software to satisfy these new computational needs. New techniques in the areas of networking, algorithm, visualization, storage and retrieval of data within very large database are also required to permit such applications as interactive analysis and control of data production in such a computing environment.

To satisfy these needs, graphical interfaces have to be designed and developed not on the basis of abstract principles but on the experimental verification of their perceptive and cognitive properties.

We assume that computer systems would be more congenial and easier to use if the comunication protocol, lexicon or graphical interface were more natural and familiar to theuser. The human factor research on the interaction between user and information system suggests that dramatic improvements of current systems are possible. In particular, a well–designed graphical interface based on knowledge of the user's cognitive ability can significantly improve the user–system interaction.

In the graphical environment, icons representing objects or functions are generally more informative and easier to manipulate than verbal labels. Icons are symbols or pictorial representations analogous to physical objects, actions or functions. Because icons should resemble the functions they represent, pre–testing should be done to determine whether the user understands what the icons are supposed to symbolize.

Humans are able to perform the same task in many ways and with many degrees of facility, depending on their state of preparation, their knowledge of facts and procedures and their internal organization to use them. Therefore, other factors should be considered in selecting symbols, such as the degree of association of the iconic symbol with the concept to be portrayed.

Experiments on all these factors could produce useful information about the transparency of the icons and could suggest that some of them have to be accordingly modified.

In the present work we aimed at establishing whether the criteria adopted to shape the icons of a graphic tool for the analysis of multidimensional field were shared by the user. After a short introduction of the tool (Sect. 2) and of the background art on user–computer interface (Sect. 3), we present the method used in the test (Sect. 4). Results and discussion are presented in Sect. 5 and Sect. 6, respectively.

2 Visualization of data

The visualization, i.e. graphics, images and presentation of information in image form, is an integral part of modern environment. For example, visualization allows viewers to perceive patterns and relationships which may be missed in huge tables of numbers. However, it should be noted that in addition to forming a three dimensional view of data, a goal is to provide tools and systems which allow the user to extract information from data (Farrell 1991).

Traditionally, visualization techniques are primarily applied to pre– or post–processing of data. For example, AVS (Upson *et al* 1989), apE (Dyer 1990) and SGI Explorer (SGI 1993) are implementations of such type of visualization application based on the data flow architecture (Haeberli 1988).

Important recent advances include the use of visualization tools to depict objects of numerical libraries (see, for instance, Gilbert *et al* 1992), to observe and steer the computation during run–time (see, for example, Boubez *et al* 1992). Furthermore, the introduction of parallel computing and its realization on varied parallel architectures has necessitated the collection of run–time data that show the performance and flow of parallel computations (see, for example, Szelényi & Zecca 1991). Graphical rapresentation of these data is the only way to perceive changes and take appropriate actions.

Implementations for pre– and post–processing of data, including the previously quoted ones, are popular because of the flexibility of mixing calculation modules with display modules, and because of their easy graphical user interface. However, data flow networks are not generally used for tracking and steering or, in general, for developing detailed algorithms. Current data flow implementations support finite sets of data structures; in order to support algorithm details they would need to support user-definable, application-specific structures.

The Balsa and Zeus system (Brown & Sedgewik 1984) provides a set of tools for designing visualization environments for algorithms. The scalar mapping technique (Hibbard, Dyer & Paul 1992) defines an infinite set of data types that can serve as types of data objects in a programming language. TraSt (Messina *et al* 1993) allows application-specific structures to be defined by the user.

MUDI3 (MUltidimensional Data Interaction in 3–dimensional space) is a prototypal system (Moltedo *et al* 1993 and references therein) which maps physical parameter types as int, real, pairs and triples of real numbers to tuples related to a graphical display as color, geometric representation, icons.

The goal of MUDI3 is to represent scalar and vectorial fields defined on a three–dimensional domain.

In this system, the visualization of data is realized through three steps. The specific parameters of each of these steps are controlled by a visual interface based mainly on icons. The first step defines the display of graphical and geometrical attributes (messengers) that the viewer decides to select for the representation of data. The second step associates the messengers to the physical parameters of the field (functions). The third step controls the function or the group of functions to display and the domain set of the field to visualize (image performance).

The system of icons available to the user for the control and management of the MUDI3 has been submitted to different tests in this work.

3 User–computer interface

One of the most successful improvements to the user–computer interface made within the last decade is the inclusion of icons (Blattner, Sumikawa & Greenberg 1989). Literally, the term icon means a highly representational image, but in computer science it has very often assumed a broader meaning including both representational image and visual symbol.

Icons are symbols or pictorial representations analogous to physical objects, actions or functions. In the graphical environment, icons representing objects or functions are generally more informative and easier to manipulate than verbal labels for two reasons, at least. First, iconic communication is successful because its imagery relies on the human ability to quickly perceive natural form and shape, and second and more trivial, icons can represent much information in a small amount of space. It had also been found that subjects could remember pictures far better than equivalent sentences, and that decision based upon pictorially mediated information could be made faster (Blankenberger & Hahn 1991).

However, because icons should resemble the functions they represent, the "articulatory distance" (i.e. the difference between a picture and its meaning) must be analysed. Then, a test should be done to determine whether the user understands what icons are supposed to symbolize. The evaluation of the articulatory distance is crucial for the man–computer interaction because the smaller the distance, the smaller should be the mental workload of the user. In other words, the quality of a problem representation influences the ease with which a problem can be solved (Hayes & Simon 1976, and Newell & Simon 1972). There is also a limit to the number of new symbols which a subject can be expected to remember. If this limit is exceeded, then the usefulness of the symbol will be reduced (Murray & McDaid 1993).

The main problem in evaluating a computer interface refers to the user's profile. In particular, the literature reports many examples of qualitative and quantitative cognitive differences between experts and novices. The most obvious difference is that experts know more about the domain than novices do, but is their knowledge structured differently? According to Murphy and Wright (1984) three different answers seem possible.

First, experts might simply have more concepts even though they share with novices the same general structure of the knowledge. In other words, experts could make much finer distinctions and use more specific categories. Thus, expert–novice differences may be related to poorly formed, qualitatively different, or nonexistent categories in the novice's representation.

Second, experts' concepts might differ structurally from those of novices. Experts could form abstract, conceptually based representations, whereas novices could form concrete syntactically based representations, and tend to retain the surface elements of the problem (Adelson 1981). For example, in a computer program domain, the representation of experts is more abstract and contains more general information about *what* the program does, whereas the representations of novices is more concrete and contains information about *how* the program functions (Adelson 1984).

Experts' abstract nature of the representations was clearly shown by Chase & Simon (1973), who demonstrated that master chess players were able to reconstruct with an accuracy greater than 90% midgame boards that they had seen for only 5 seconds. The result suggested that master chess players recalled clusters instead of single units, and that they saw the single chess pieces as integral parts of larger, meaningful units. Thus it appears that experts' superiority in memorizing chess board positions arises from the existence of a large store of intact and well–organized chess configurations or patterns in memory. That is, the concepts of experts may be said to be more tighly clustered.

The same difference in representation is found examining how experts and novices categorize problems in physics: novices tend to categorize according to surface structures, whereas experts categorize according to the abstract physics principles which can be used to solve problems (Chi, Feltovich & Glaser 1981)

Third and more surprising, experts could have the same number of categories, each of them with the same structure as novices', although experts know more about the objects in their domain of expertise. Then, according to this last possibility, the hypotesis of no difference between experts and novices is a realistic one, at least in some domain.

Because different user levels are possible between novices and experts in using MUDI3, we selected three groups of subjects for our experiment: novices, intermediates and experts.

4 The experiment method

The design of the experiment aimed to evaluate the level of transparency of a set of icons of MUDI3 and the cognitive congruity between icon and its definition and label. Therefore we defined subjects, stimuli and procedure as follow.

4.1 Subjects

Subjects were selected among students and researchers of the university of Bologna. All of them were naive as to the purpose of the test. Three groups were formed.

The first group (novices) was made up of ten students of literature who admittedly had no previous experience in human–computer interaction.

The second group (intermediates) was made up of ten student of science courses; they had a good knowledge of the background principle of human–computer interaction and had up to five year of experience with interactive tools.

The third group (experts) was made up of ten researchers in science or computer science and had more than five years of experience with computers and interactive tools.

4.2 Stimuli

Stimuli were the 24 colored icons (see Figure 1) used in MUDI3 to represent different actions and functions; 28 definitions, 24 linked to the icons and 4 without any correspondence to the icons of Figure 1 but describing standard functions for visualization; and 28 labels, 24 defining the icons and 4 referred to other concepts typical of graphics tools.

4.3 Procedure

All the subjects of the three groups were asked to answer questions of three tests.

In the first test, each subject received a colored picture with the icons shown in Figure 1 and labeled with numbers (for instance, ic 2), and a sheet of paper with definitions of different length labeled with letters (for instance, *S: The background color can be changed*). The task consisted in the association of a definition to an icon. The subjects were told that there were four extra definitions and that they could associate more definitions to the same icon. It was stressed that the task was not under time pressure.

In the second test, each subject received the icons' picture shown in Figure 1 and a paper with a list of labels consisting at the most of four words and marked with letters (for instance, *K: rotation*). The task consisted in the association of a label to an icon. The subjects were told that there were four extra labels and that they could associate more labels to the same icon. Also in this second test it was stressed that the task was not under time pressure.

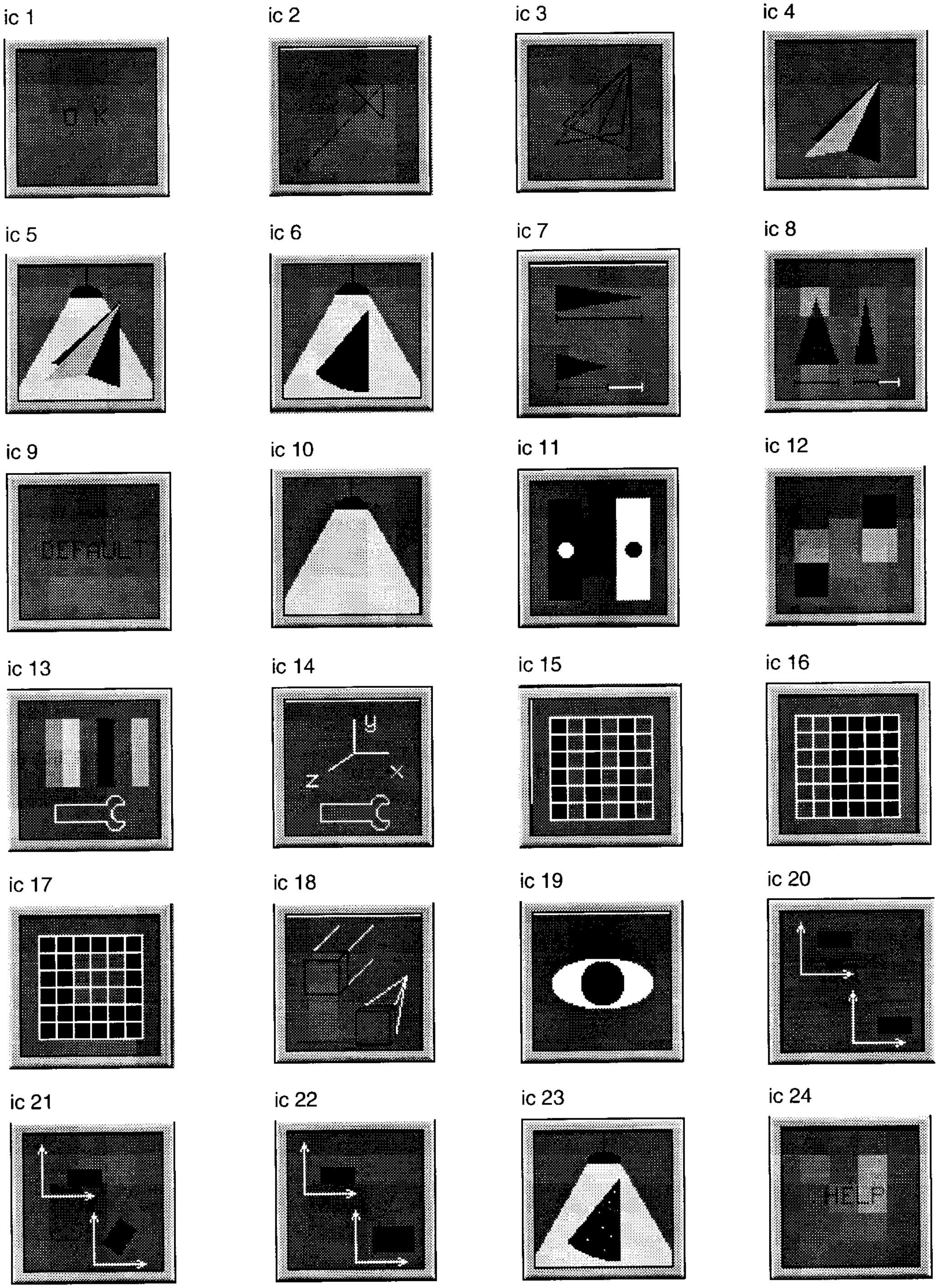
In the third test, our goal was the evaluation of the cognitive compatibility between the icons and the concepts that they are supposed to symbolize. The subjects were faced with the picture of the 24 icons matched with the 24 correct labels indicating the functions symbolized. Each subject had to judge the articulatory distance between each icon and its meaning employing a five–point evaluation scale. The five possible evaluations were: i) totally inadequate; ii) partially inadequate; iii) adequate; iv) very adequate; v) totally adequate.

5 Results

The number of errors observed in the test 1 and 2 were entered in two one–way analyses of variance (one for test 1 and one for test 2). The between-subjects factor was Experimental group with three levels: novices, intermediates and experts.

When the subjects had to match the icons with the definitions, the analysis of variance showed a significant effect of the three–level factor: $F(2,27) = 9.30, p < 0.0008$. The number of errors were 107, 42 and 52 for novices, intermediates and experts, respectively, corresponding to 44.5%, 17.5% and 21.6% of the total of the responses.

A comparison among the three levels of the factor using the Newman–Keuls test, showed that novices made more errors than intermediate and expert subjects ($p < 0.01$), but no difference arose between these two groups.



ic 1 ic 2 ic 3 ic 4

ic 5 ic 6 ic 7 ic 8

ic 9 ic 10 ic 11 ic 12

ic 13 ic 14 ic 15 ic 16

ic 17 ic 18 ic 19 ic 20

ic 21 ic 22 ic 23 ic 24

Figure 1: The icons of MUDI3 whose cognitive properties have been tested

Table 1: Results of test 3

Evaluation	Novices	Intermediates	Experts
totally inadequate	04.5%	04.5%	02.0%
partially inadequate	08.5%	10.0%	08.3%
adequate	13.3%	11.6%	20.0%
very adequate	25.8%	27.5%	26.4%
totally adequate	47.9%	46.2%	42.9%

Also in the task of matching the icons with the labels, the analysis of variance showed a significant main effect: $F(2,27) = 4.87, p < 0.015$. The number of errors were 75, 32 and 41 corresponding to 31.2%, 13.3% and 17.0% for novices, intermediates and experts, respectively. A Newman–Keuls test confirmed the pattern of results obtained in test 1: the novices made more errors than the other two groups ($p < 0.05$), but the intermediates and the experts were not significantly different.

The Table 1 reports results obtained in the task of evaluating the cognitive compatibility between the icons and the concepts that they are supposed to symbolize.

6 Discussion

The results of test 1 and 2 are partly expected and partly unexpected. On the one hand, in fact, data show that the group of novices do more errors than the other two groups in agreement with the role played by experience and training on computer users.

On the other hand, experts make more errors than intermediates in both tests (21.6% *vs.* 17.5% and 17.0% *vs.* 13.3%, respectively). Although these differences are not statistically significant, the result represents an interesting trend to discuss. The introduction of the intermediate group, which was not investigated in previous works, offers some suggestion concerning the cognitive differences for different levels of expertise. One could argue that intermediates and experts share the general problem representation with experts having more concepts, especially more specific concepts to deal with the greater number of details they know.

In our opinion this is reflected in the fact that the experts have matched more definitions or labels with one only icon and viceversa, with a consequent increase of errors. Therefore, the greater knowledge domain of the experts could justify both better performance with respect to the novices, and worse performance with respect to the intermediates.

The second interesting point comes from the different results obtained in test 1 and 2. A new analysis of variance showed a significant difference between errors in test 1 and 2: $F(1,27) = 14.65, p < 0.0007$. The pattern of results (201 errors in test 1, and 148 errors in test 2, overall) demonstrates that the subjects judged more compatible the labels than the definitions of the icons. Therefore, many definition have to be modified or newly written.

In test 3, the evaluation of the novices and the intermediates are very similar. Instead, the experts expressed the judgement *adequate* more frequently (20.0% *vs.* 13.3% for novices and 11.6% for intermediates), whereas the judgement *totally adequate* was reported less frequently (42.9% *vs.* 47.9% for the novices and 46.2% for intermediates). This is in agreement with the structure of knowledge of experts as an integrated conceptual model, according to which it is easier to judge vaguely *adequate* an icon, but more difficult to admit that an icon is *totally adequate.*

Although the goal of the present work has been a general evaluation of icon system of MUDI3, the data permit some preliminary observation on single icons. For example, summing errors of test 1 and 2, icons 0, 9 and 15 realized the worst performance scoring 41.5%, 49.9% and 58.8%, respectively. Whereas the best performances in term of minor overall percentage of error were obtained by icons 10 and 20 (10.0%), icons 11 and 17 (11.6%), and icons 8 and 23 (13.3%).

Research is under way to analyze deeply results on each icon. New data should permit, if necessary, a redesigning of icons in order to obtain a better cognitive compatibility between icons and users.

Acknowledgements

This work has been partially supported by C.N.R., Progetto Finalizzato Sistemi Informatici e Calcolo Parallelo and by European Union contract No. CHRX–CT93–0410 (Human Capital programme).

References

[1] Adelson, B.: 1981, *Memory & Cognition* **9**, 422

[2] Adelson, B: 1984, *J. of Exp. Psych.: Learning, Memory and Cognition* **10**, 483

[3] Blankenberger, S., Hahn, K.: 1991, *Int. J. of Man–Machine Studies* **35**, 363

[4] Blattner, M. M., Sumikawa, D. A., Greenberg, R. M.: 1989, *Hum. Comp. Interac.* **4**, 11

[5] Boubez, T. I., Froncioni, A.M., Peskin, R. L.: 1992, *ACM Trans. Math. Softw.* **18**, 1

[6] Brown, M., Sedgewick, R.: 1984, *Comp. Graphics* **18**, 177

[7] Chase, W. G., Simon, H. A.: 1973, *Cognitive Psyc.* **4**, 55

[8] Chi, M. T. H., Feltovich, P., Glaser, R.: 1981, *Cognitive Science* **5**, 121

[9] Dyer, D.: 1990, *Comp. Graphics and Appl.* **10**, 60

[10] Farrell, E. J.: 1991, *IBM J. Res. Dev.* **35**, preface

[11] Gilbert, J. R., Moler, C., Schreiber, R.: 1992, *SIAM J. Matrix Anal. Appl.* **13**, 333

[12] Haeberli, P.: 1988, *Comp. Graphics* **22**, 103

[13] Hayes, J. R., Simon, H. A.: 1976, *Cognitive Psyc.* **8**, 165

[14] Hibbard, W., Dyer, C., Paul, B.: 1992, Preprints, Conf. Interactive Information and Processing Systems for Meteorology, Oceanography and Hydrology. Atalanta, American Meteorology Society, 101

[15] Messina, A., Pensato, M., Borghesi, M.: C.N.R. Technical Report on Project *Sistemi Informatici e Calcolo Parallelo*, n. 1/147, March 1993

[16] Moltedo, L., Ascani, F., Messina, A.: 1993, *NATO ASI Series* **106**, 203

[17] Murphy, G. L., Wright, J. K.: 1984, *J. of Exp. Psych.: Learning, Memory and Cognition* **10**, 144

[18] Murray, B. S., McDaid, E.: 1993, *Int. J. of Man–Machine Studies* **38**, 23

[19] Newell, A., Simon, H. A.: 1972, *Human Problem Solving* NJ, Prentice–Hall

[20] SGI: *IRIS Explorer manuals* Silicon Graphics, 1993

[21] Szelényi, F., Zecca, V.: 1991, *IBM J. Res. Dev.* **35**, 270

[22] Upson, C., Faulhaber, T. Jr., Kamins, D., Laidlaw, D., Schlegel, D., Vroom, J., Gurwitz, R., van Dam, A.: 1989, *Comp. Graphics and Appl.* **9**, 30