

THE USE OF GUIDELINES IN MENU INTERFACE DESIGN: Evaluation of a draft standard

Flavio de Souza*
University College London
Ergonomics Unit, 26 Bedford Way
London, WC1H 0AP, England

Nigel Bevan
National Physical Laboratory
DITC, Teddington, Middlesex
TW11 0LW, England

This paper reports a case study of the evaluation of the effectiveness of a draft standard containing human factors guidelines for menu interface design. Three designers were given a week to study the guidelines, before spending a day using the guidelines to redesign a menu interface. They were asked to justify their design in terms of the guidelines. The designers made errors or had difficulties with 91% of the guidelines. The cause of the errors and difficulties was analysed, which enabled recommendations for improvements to be made. Despite the difficulties with interpretation of the guidelines, the resulting interfaces only violated an average of 11% of the guidelines which could be assessed. It is concluded that it was difficult for the designers to integrate detailed design guidelines with their existing experience.

INTRODUCTION

In the 1980's the point was reached where the downward spiral in cost of computer technology made it more cost-effective for computers to serve the needs of people, than for people to serve the needs of computers as had previously been the case (Bevan and Murray, 1985). To create systems which are subservient to human needs requires a change of priorities in the design process. Ideally this means that systems designers should also be experts in human factors. However, not only are such multi-disciplinary individuals very rare, but there are not even enough trained experts in HCI to participate in every design team.

There is thus a great demand for mechanisms to transfer human factors knowledge to existing designers (Hannigan and Herring, 1987). One widely-used means is collections of design guidelines. These vary in their degree of specificity, ranging from those which state a small number of general principles (eg DIN, 1987), through those giving detailed guidance derived from cognitive psychology (eg Marshall, Nelson and Gardiner, 1987), to specific advice based on empirical evidence (eg Smith and Mosier, 1984b).

Guidelines can be used in a number of different ways. They are a useful compilation of HCI knowledge, and can provide the human factors professional with an authoritative source of advice for designers. They can also be used as a means of transferring knowledge to designers as part of educational or training courses. But the most valuable role they could play would be as a direct source of reference and guidance for designers to use during the design process. However, surveys have shown that in practice only a minority of designers consult such guidelines (Smith and Mosier, 1984a, Hammond et al, 1983).

* now at Birkbeck College, Psychology Department, Malet Street, London, WC1E 7HX, England.

The ISO Software Ergonomics group (Holdaway and Bevan, 1989) in conjunction with the

American National Standards Institute has been developing a set of guidelines for menu interface design (Williams, 1989), incorporating material from the MITRE Corporation guidelines (Smith and Mosier, 1984b). It is shortly intended to publish these guidelines as the first draft of an ISO standard. The objectives of the study reported here were to evaluate the extent to which designers are able to use such guidelines to design menu interfaces, and to suggest how the guidelines could be improved.

METHOD

A task such as using guidelines to design a menu interface can be analysed in terms of the reduction of the 'designer costs' (cognitive and affective) when using guidelines, and the improvement of the 'quality of the product' (in this case the designed interface) (Dowell and Long, 1989). For the purpose of this study 'cognitive costs' are expressed in terms of errors (recognized and unrecognized) and difficulties in using the guidelines, and 'affective costs' refer to the feelings of distress (irritation, annoyance, anger) experienced by the designers using the guidelines. The 'product quality' can be assessed by comparing the the initial and redesigned interfaces.

It should be emphasised that what is being tested here is the extent to which the guidelines successfully transfer human factors knowledge, and not the correctness of the scientific findings or principles that constitute their basis.

Analytic Structure

The evaluation is based on an extension of the formal-empirical approach to analysis of user behaviour proposed by Long and Buckley (1987). The analytic models concerned have been described and illustrated in de Souza, Long and Bevan (1990). The source of errors or difficulties can be divided into four categories.

1. Interpretation of the guidelines.
2. Identification of the design problem.
3. Solutions to the design problem.
4. Implementation of the solutions.

In order to interpret the designers' behaviour, an initial 'Model of the Sources of Errors and Difficulties' was constructed (see Figure 1) which identified problems which could be due to lack of information, either about the procedure to follow to apply a guideline, or about the design goals and benefits; or lack of clarity of information, either about the conditions under which the guideline should be applied, or about the precise nature of the proposed solution.

In this analysis, *errors* are design decisions which contravene a guideline. The errors may be *recognized* or *unrecognized* by the designers. *Difficulties* are design decisions which are consistent with the guidelines, but which have a high cognitive or affective cost for the designer (eg in terms of work load, effort or frustration).

The Experimental Design

Three interface designers used the guidelines to improve the design of the structure and presentation of the menu options of a badly designed traditional menu-style interface for an electronic mail system. The technique employed for the assessment of designers' cognitive and affective costs was an interventional observational study. The experimental sessions were video-recorded, and the errors and difficulties in using the guidelines were identified by means of

protocol analysis and observations made by the experimenter.

The Subjects

The three subjects employed in the experiment were designers with a minimum of 1 year of experience in designing interactive programs.

The Guidelines

The guidelines used were those drafted for the International Standards Organisation (Williams, 1989) and consist of 87 guidelines. However, because of time constraints, the present experiment addressed only those guidelines concerned with 'design requirements for menu structure' and 'design requirement for menu presentation', which constituted a total of 45 guidelines.

Preparation of the Subjects

Seven days before the experiment the designers received the set of guidelines, and general instructions for the experiment, which informed them about the objective, methodology and scheduling of the experiment. They were asked to read the guidelines very carefully and to mark the parts that they found unclear, and to make notes about any difficulties (particularly about the interpretation of the guidelines and their applicability to real interface problems).

The Task

The designers were instructed to improve the design of the structure and presentation of the menu options of the initial interface. In order to do this they were required to use and apply the menu interface guidelines.

The initial version of the interface design was presented to the designers on paper and they were asked to redesign it by drawing on a white board. In order to make the task more realistic the designers were given the objective of designing for a real-world problem, which included relevant information about the characteristics and objectives of the organization which would use the system, including the characteristics of the users and their tasks.

Procedure

1 Study of the set of guidelines

During the preparatory part of the experiment the designers were asked to study the guidelines for 45 minutes. This was to ensure that they had sufficient knowledge of the guidelines and to serve as a warm up period for the actual experiment.

2 Familiarisation with the initial interface

The second phase of the experimental session consisted of presenting the initial version of the interface to the designers. The designers were asked to read the contents of the menu panels, the description of the functionality and the instructions about how the system worked.

3 Use of the guidelines in the experimental session

It was observed during a pilot study that designers tended to identify and solve the design problems by an iterative and random procedure. Even a designer who was emphatically asked to use the guidelines tended to put them to one side and rely on his existing experience and prior reading of the guidelines. When asked from time to time to justify his implemented design solutions by referring to the relevant guidelines, the designer appeared to become confused, unable to find where most of the supposed justifications were in the guidelines. It was therefore decided to divide the design process into two distinct parts.

Part I

The experimenter clarified the nature of the task, and explained how the initial menu interface worked. During this session the designers were oriented towards changing every feature of the interface design that they thought should be immediately changed, based both on their previous study of the guidelines and on their own experience. They were not obliged to follow the guidelines but they were asked to think aloud and justify every design decision that they made without necessarily showing the specific guideline that they were applying. The designers had one hour to check and design the main features of the new version of the interface.

Part II

During this phase (lasting three hours) the designers were asked to change the new interface to make it consistent with the guidelines by using and applying all the guidelines systematically one by one. The designers were asked to apply each guideline to 'identify' and 'solve' design problems in the new version of the interface (the product of the design process part I). During this part of the design process (the most relevant to the experiment) the experimenter asked the designers to clarify what they were trying to do (intended design solution), why they were trying to do it (identified design problem) and indicate which guideline they were using throughout the design process. It was made clear to the designers that the intervention was not intended to change the designers' behaviour.

Application of the guideline either led to confirmation of an already identified problem and implemented design solution (in either version of the interface), or to the identification of a design problem and the implementation of a design solution proposed by the guidelines, changing the feature of the current interface under consideration.

The interview

After the design task the designers were interviewed by the experimenter for 45 minutes. The objectives of the semi-structured interview were: (a) to collect complementary information to support the assessment of the affective and cognitive costs to the designer, (b) to discuss the doubts, comments, criticisms and suggestions made by the designers before and during the experimental session, to help the experimenter to identify and understand any sort of difficulties and errors that may not have been observed during the experiment and (c) to discuss and clarify, if necessary, any difficulties and errors observed by the experimenter during the experimental session.

RESULTS

Evaluation of Effectiveness

Cognitive Costs

On average each designer had errors or difficulties with 66% of the guidelines, and 91% of the guidelines produced errors or difficulties for at least one designer. The percentage of guidelines which showed at least one error or difficulty in each category was:

1. interpretation: 44% (none were recognized);
2. identification: 47% (including 2% which were recognized);
3. solution: 31% (including 18% which were recognized).

No errors or difficulties of implementation were observed. This is probably the result of the simplified simulation of the menu design task without on-line implementation.

Examples of the errors and difficulties are given below. In some cases a designer made more than one category of error on a guideline.

Affective Costs

The observed affective costs of the guidelines were low: no evidence was detected in the protocol analysis or interviews made by the experimenter.

Quality of the Product

The degree of inconsistency of each designed interface with the guidelines was expressed as the percentage of guidelines violated in relation to the total number of guidelines which could be assessed (32). (The remaining 13 guidelines were either not applicable to the interface design process or to the quality evaluation process.) The assessment of product quality was made by expert analysis of the extent to which the final interface conformed with the guidelines (irrespective of the encountered errors or difficulties).

As can be seen in Table 1, on average only 11% of the guidelines were violated in the final design. In many cases it was possible to produce a satisfactory design despite the difficulties experienced in applying the guidelines. (However, there are other aspects of the quality of the design not represented by this particular assessment which may also be related to the number of errors and difficulties.)

Types of Problem and Design Recommendations

This section illustrates with examples the different types of problem observed for each category or error, explaining the cause and making appropriate design recommendations. The same types of problem could also cause difficulties, where the correct solution of a design problem was preceded by difficulty of interpretation of a guideline.

1. Interpretation of the Menu Guidelines

Example 1.1. Lack of clarity of the conditions under which the guideline should be applied, causing an unrecognized error of interpretation of the guidelines

Guideline 5: 'Natural grouping: group options by function or into logical categories which are meaningful to users'

Comment by Designer: "To some extent it is repeating what has already been referred by the guideline 1."

Cause: the textual structure. This guideline is about groups within a menu panel while Guideline 1 is about the design of menu panels (group the options into menu panels).

Recommendation: structure the text of the guideline so as to prevent the designer from misinterpreting the condition under which the guideline should be applied. For example, reword using italics, as follows: '*In designing groups within a menu panel*, group options by function or into logical categories which are meaningful to users.'

When the designers went through the guidelines they tended to forget that guidelines were listed under headings (in this case: 9.2.4.1.B Groups within a menu panel). Another reason why this designer interpreted guidelines 1 and 5 as the same may have been because both of them are based

on the same principle and have similar headings: ‘Natural groups’ and ‘Natural grouping’.

Example 1.2. Lack of clarity of the conditions under which the guideline should be applied, causing a difficulty in interpretation of the guideline.

Guideline 12: ‘Distinctive designators: if designators are used (a) the user should be able to clearly distinguish the designator for each option, and (b) do not mix upper and lower case codes for menu option designators.’

Comment by Designer: “I ignored that because I did not know what ‘designators’ meant”

Cause: the word ‘designator’ was not understood..

Recommendation: replace ‘designators’ with another word better known to designers, eg ‘labels’.

Example 1.3: Lack of clarity of the nature of the proposed design solution, causing an unrecognized error of interpretation of the guidelines.

Guideline 25: ‘Horizontal space columns: columns of options should be separated, horizontally, by at least three character spaces.’

Comment by Designer: “I have applied it in all the menu panels by putting a gap of three spaces between the labels and the option names.”

Cause: lack of clarity of the term ‘columns of options’ which is intended to include option names and associated labels.

Recommendations: provide clarification of the term ‘columns of options’, eg by a note such as: ‘*column of options* refers to the option names and respective labels’. It may also reduce the cognitive costs of the designers if the implicit conditions to apply this guideline are made explicit, eg: ‘*If a menu panel has more than one column of options* the columns of options should be separated, horizontally, by at least three character spaces.’

2. Identification of Menu Design Problems

Example 2.1. Lack of information about the procedure to apply a guideline, causing an unrecognized error of design problem identification.

Guideline 10: ‘Frequency ordering: If the frequency of options use is known and option groups are small (eight or less), place the most frequently used options first.’

Comment by Designer: “Frequently used will be ‘redisplay’ and ‘logout’. On the basis of the available information this is the best that can be done, no information is available about it. Based on the lack of information I have to use what I call an intuitive order.”

Asked about the specific information that was not available he said: "I need information like, for example: if you have only one list you would be more likely to use this list more frequently. ‘Display list’ option in the ‘Mail list menu’ could be *there*, but I do not know, it is a bit vague. I need to know what is exactly in there in order to infer the most frequently used options. Not having the exact information about what is on the list I cannot think about a better ordering.”

Cause: lack of information about the procedure to apply and assess frequency of options use. When the designer referred to the need for exact information he does not mean empirical

information about frequency of use by the users. He refers to the need to have a well-structured and explicit system in order to make a logical inference about frequency of options use.

Recommendation: provide a procedure for how to assess 'frequency of options use'. The designer inferred that 'logout' will be frequently used because the users will have to use it every time that they logon to the system. However this does not necessarily mean that it is going to be frequently used. This incorrect inference reinforces the need for a very clear guideline.

Example 2.2. Lack of information about the procedure to apply a guideline, causing a recognized error (a design decision inconsistent with the guidelines).

Guideline 1: 'Natural groups: If options can be arranged into natural or conventional groups known to users, organize the options in levels and groups consistent with the natural hierarchy.'

Comment by Designer: "This guideline would be applicable for partition of options into menus. I am confused: does it apply to options within the menus? It would be more applicable to group the options into menu panels." He went on to apply the 'natural grouping' prescriptions to organise menus within panels, a design problem not addressed by the guideline.

Recommendation: The scope of the guideline should be clarified.

3. Solution of Menu Design Problems

Example 3.1. Lack of information about the procedure to apply a guideline, causing an unrecognized error of design problem solution.

Guideline 39: 'Familiar terminology: use terminology familiar to users.'

Comment by Designer: "There is no standard person, so what is familiar to me is not necessarily familiar to the users. Some of the users might have some knowledge of computers as a hobby and others not. That is why I have tried to use simple words."

Cause: Lack of information about the procedure to apply the guideline and assess familiar terminology. Even though this guideline explicitly demands consideration of the knowledge of the end user, the designer relied on his own intuition or common sense on which to base the design decision.

Recommendation: provide information about the procedure to apply the guideline and assess familiar terminology. For example: 'Note: make sure that the terminology used is familiar to users by carrying out an empirical test using a sample of the target population.'

Example 3.2. Lack of information about the design goals and benefits, causing a recognized error of design problem solution.

Guideline 45: 'Branching to submenus: if an option leads to another menu rather than to executing an action, then provide appropriate cues to the user. Example: options leading to a submenu could be followed by a right-pointing arrow at the end of the option label or "menu" might be included in the option name.'

Comment by Designer: "I would not implement this guideline in this interface because it does not add anything. It is of no utility at all to this interface, because so long as you select an option you

change the thing that you were displaying. I think that it would be useful for menus that use a mouse because you would be able to move the cursor to the 'arrow', click on it and another menu would appear on that point in particular.”

Cause: lack of information about the design goals and benefits. It seems that this designer has misinterpreted the guidelines based on previous experience with pop-up menus. He may not have experienced this guideline implemented on a traditional menu interface, and so did not envisage the utility (design goal) of implementing it.

Recommendations: provide information to the designers making explicit that the design goal is to improve perceptual distinctiveness among the option names.

CONCLUSIONS

It can be seen from the detailed analysis of the errors and difficulties that if guidelines are to be correctly interpreted, it is essential that they contain clear information about:

- the design goals and benefits
- the conditions under which the guideline should be applied
- the precise nature of the proposed solution
- any procedure which must be followed to apply the guideline

To produce effective guidelines requires careful attention to these design principles for guidelines, complemented by an empirical evaluation such as the one described in this paper. The finding that only 4 guidelines out of 45 did not produce any problems for the three designers illustrates the importance of designing a set of guidelines which are not only accurate, but also usable. This is potentially a problem when publishing guidelines as a standard, as existing rules for producing standards (eg ISO, 1989) which are concerned with technical accuracy, make it difficult to present the information in a manner which optimizes usability.

With iterative development of the guidelines it should be possible to rectify the errors and difficulties of interpretation, identification and solution. The recognized errors are the most difficult to deal with, as they suggest that designers may apply what they understand to be the principles underlying the guidelines, to design problems for which they were not intended. In other words, designers are not passive consumers of guidelines advice, but actively seek to interpret them in the light of their own experience. This makes it particularly important to clarify the design goals and objectives, and the procedures which may be used to achieve these objectives. This will reduce the cognitive costs to the designers and decrease the likelihood of the guidelines being misinterpreted.

Iterative refinement will maximise the potential usefulness of the guidelines, but it cannot itself guarantee that designers will actually use them. The experience in this study was that designers were very reluctant to make explicit reference to individual guidelines, except when directly asked to do so. The attitude is summarised by one designer:

“I see these guidelines as an explicit statement of my intuitive knowledge. I had already got this information but I had not perceived it. So, I would know what is wrong and I would have an idea about what is right on the interface but I would not be able to explain these perceived ideas. ... I do not need these guidelines because I know how to do it [design interfaces] already. I think that it is important to read it one or two times and once you have got this information from the back of your mind to the front of your mind, you do not need it anymore, so you put it away. I think that you will forget about it. Possibly, if you have a problem you might have a quick look at it.”

The results from these designers support the view (Taylor and Galer, 1989) that the most effective way to ensure that human factors are addressed is to integrate the human factors issues into the normal design process. Guidelines alone do not appear to be a very effective way of achieving this.

The guidelines are likely to have more impact when they are published as a standard. The ISO group has yet to decide whether it is possible to measure conformance with a guidelines standard. If a solution is found it will provide a powerful incentive for designers to carefully evaluate the design for a menu interface against the individual guidelines.

More generally, it may be useful to establish instructional workshops to give designers the opportunity to integrate the principle in the guidelines with their existing expertise.

ACKNOWLEDGEMENTS

This paper is based on research carried out by the first author in conjunction with the National Physical Laboratory in part fulfillment of the degree of MSc. (Ergonomics), University of London and was funded by Coordenadoria de Aperfeiçoamento de Pessoal de Nível Superior and Universidade Federal Fluminense (Brazil).

The authors would like to acknowledge the contributions made by Prof. John Long.

REFERENCES

- Bevan N. and Murray D. (1985). *Man-Machine Integration. Infotech State of the Art Report, 13*, 1, (Maidenhead: Pergamon Infotech Ltd).
- DIN (1987) *DIN 66234 Part 8: Principles of dialogue design*.
- Dowell J. and Long J.B. (1989). Towards an engineering conception for the human factors discipline. *Ergonomics*, (in press).
- Hammond, N.V., Jørgensen, A.H., MacLean, A., Barnard, P.J. and Long, J.B. (1983). Design practice and interface usability: evidence from interviews with designers. In *Proceedings of CHI'83*, Boston. (New York: ACM.)
- Hannigan S. and Herring V. (1987). Human factors in office product design: European practice. In: *Cognitive engineering in the design of HCI and expert systems*, ed. G. Salvendy (Amsterdam: Elsevier).
- Holdaway K. and Bevan N. (1989). User system interaction standards. *Computer Communications*, **12**, 2, 97-102.
- ISO (1989). *Directives Part 3: Drafting and presentation of international standards*. (Geneva: International Standards Organisation).
- Long J.B. and Buckley P. (1987). Cognitive optimization of videotex dialogues: a formal-empirical approach. In *Psychological Issues of Human Computer Interaction in the Work Place*. eds. M. Frese, E.Ulich, W. Dzida (Amsterdam: Elsevier).
- Marshall C., Nelson C. and Gardiner M.M. (1987). Design guidelines. In *Applying Cognitive Psychology to User-Interface Design*, eds. M. M. Gardiner and B. Christie (Chichester: Wiley & Sons Ltd).
- Smith S.L. and Mosier J.N. (1984a). The user interface to computer-based information systems: a survey of current software design practice. *Behaviour and Information Technology*, **3**, 195-203.
- Smith S.L. and Mosier J.N. (1984b). *Design guidelines for user-system interface software*. Technical Report NTIS No. AD A154 907.
- de Souza, F.L. (1989). Ergonomic evaluation of menu interface guidelines. (National Physical Laboratory Technical Memo).

- de Souza F.L., Long J.B. and Bevan N. (1990). Types of error and difficulty in using human-factors guidelines. In *Contemporary Ergonomics, 1990*. (London: Taylor & Francis).
- Taylor B. and Galer M. (1989). User-centred design practice in office automation. In: *Proc. HCI International 89*.
- Williams J.R. (1989) *Menu Design Guidelines*. ISO TC159/SC4/WG5 WD 9241-14 (17 March 1989).