Now, where was I?

GRAEME FIELD and ROBERT SPENCE

Consultant Psychologist
New Zealand

Imperial College
London

Abstract

Although a very common everyday occurrence, the disruptive effects of interruptions on users of computer systems has thus far received little attention. In the experiment reported here, interruptions were found to have a significant effect ($p=0.001$) on the ability of subjects to answer questions by reference to a tree-structured, menu-accessed data base, this ability being reflected in search time and in the number of frames accessed. On average, search time increased by 46% and the number of frames accessed increased by 25% through interruption. Directions for future research are suggested.

Interruptions

Real practical use of information systems is frequently characterised by interruptions. An incoming telephone call will temporarily halt the search for some data in an information system, or a colleague's remark may interfere with the preparation of a document on a word processor. With notable exceptions (Gillie & Broadbent, 1989; Kreifeldt & McCarthy, 1981), few studies have been carried out into the effect of interruptions on the use of an information system. Gillie & Broadbent (1989) found that two factors - requiring subjects to pay immediate attention to the interruption, and the complexity of that interruption - contributed significantly to the disruptive effects of the interruption. The mechanism that permits subjects to delay paying attention to the interruption and rehearse their database position remains unclear. In some respect it appears intuitively obvious that an interruption is not an interruption *unless* it requires immediate attention. In their earlier experiment, Kreifeldt & McCarthy (1981) compared Reverse Polish and Algebraic Notation calculator types rather than computer-based data base systems. Within the restrictions of their experimental approach and techniques they were able to note an increase in the amount of processing that occurred in the post-interruption period.

Within the more general context of psychological research, investigation of the effects of interruptions has focused on the interruption of rehearsal in the immediate post-learning period and prior to recall (e.g., Peterson & Peterson, 1959). There are some significant differences between these types of experiments and the one reported here. First, the Peterson & Peterson type of study presupposes that learning has taken place prior to attempted recall. In our case, although we assume that some learning is taking place, it is the effect of an interruption *en route* to the solution to a problem that is the important factor. Second, we assume that, once a subject has retrieved the answer to the current problem, that specific piece of information no longer has any particular relevance, either in recall or in solving future problems, and makes no contribution, therefore, to the ability of the subject to deal with any interruption.

An earlier study of the effects of interruptions (Field, 1987) involved 12 subjects (11 male, 1 female, average age 24.6 years) all of whom had had extensive computer experience. Two conditions were employed. In the Selective Retreat (SR) condition, the top of each frame (Figure 1) lists the previous (superordinate) selections, both to give the user information about his or her location within the data base as well as to allow retreat to any previous selection.

In the Standard Menu (SM) condition no such information is given and the user can retreat only to the immediately previous frame or to the root node. Each subject performed in either the Selective Retreat or the Standard Menu condition and attempted to answer five Simple Questions (of which two were interrupted) and five Complex Questions (of which two were interrupted) by accessing a menu-based information system. An interruption consisted of the replacement of the subject-selected frame by an 'interruption frame' containing a question which the subject was required to answer before proceeding further.

The results of this study showed that interruptions had a significant effect on both the number of frames accessed to obtain a solution and the time taken to obtain that solution. Subjects with the Selective Retreat facility required significantly fewer post-interruption frame accesses to answer the questions than did subjects in the Standard Menu group ($p=0.003$), regardless of whether Simple or Complex questions were being answered. All subjects made significantly fewer

frame accesses in the post-interruption period to complete their answers to Simple (as opposed to Complex) questions ($p=0.07$), regardless of whether they were performing in the Standard Menu or Selective Retreat condition.

This experiment was primarily an exploratory pilot study designed to discover if there was an appropriate experimental approach that could be used to study the effects of interruptions.

Some of the problems that were noted during this initial study included the small number of subjects, the type of interruption frame that was used, the necessity to reclassify question complexity, and so on. The result of addressing these issues is the study we report here.

Issues

In addition to the need to establish whether interruptions do have an effect upon data base navigation, two other factors were thought to be potentially important to the success with which information could be located in the face of interruptions. The first was question complexity. Question complexity was initially assessed by noting the number of different menu selections made at the first level (root) of the tree, for the same question, by a selection of subjects. The second related to any assistance that might be derived from the use of navigational aids. There are many available navigational aids (Heppe et al, 1985). However, the navigational aid selected for inclusion within the present study (and also within Field's earlier experiment) was the selective retreat facility already discussed and illustrated in Figure 1.

That experiments to assess the effects of interruptions should be carried out can be argued on at least two premises. First, we would wish to design future information systems in such a way that their use is robust with respect to interruptions. Second, by studying a new facet of information system use, our model of their navigation by users might be improved.

This paper, then, describes an experiment whose objective was to provide initial insight into the effects of interruptions on menu-driven information system navigation.

Method

The Data Base

The data base employed in the experiment was a videotex-style data base established on a BBC Model B microcomputer (Field & Apperley, 1990). This data base contains more than 1200 pages of information (related to entertainment, business, travel and accommodation and the like) for a fictitious city called Carlton. The root menu (containing six choices) and a typical intermediate menu are shown in Figure 2.

The information could be accessed via a hierarchy of menu frames after the style of PRESTEL (Woolfe, 1980). The majority of information pages were located at the fourth or fifth level of this hierarchy, and conventional classification practices were used in providing the index to the information

The Underground 624

11 Travel
12 Local Services
13 Rail

1 Timetables

2 Temporary closures

3 Fares and Concessions

Enter choice: _____

Figure 1 Frame corresponding to the Selective Retreat condition

very rapid removal ensured, as far as possible, that the subject would be unlikely to attempt rehearsal and would therefore give immediate attention to the interruption, a factor that will be seen to be important when we later compare our results with those of Gillie & Broadbent (1989).

Subjects were told to press the yellow key on the keypad when they had completed their interruption task, whereupon the system returned to the frame last selected by the subject. In all cases, interruption occurred at the third level from the root of the menu.

Dependent Variables

The variable **Frames Accessed** was defined as the total number of keyed menu selections made by a subject as they answered any particular question. Frames Accessed includes the initial selection from the main menu and the final selection to obtain the target frame containing the information required to answer the question. It also includes any frames accessed in the course of any retreats, whether made through the selective retreat facility or simply by a return to the main menu.

Not included in the Frames Accessed score were any selections made in the process of returning to the main menu following the successful answer to a question, or any selection that was made in the 0.1 second period between the selection of the interrupted frame and the appearance of the interruption frame.

Search Time was defined as the time elapsing from the appearance of the main menu to the selection of the correct target frame, less the time involved in dealing with any interruption. The time taken to return to the main menu following the access of the correct target frame was not included in the score. Thus, Search Time was the time that subjects were *actively engaged* in the search for an item of information.

The **Post-Interruption Pause** was defined as the time that elapsed from completion of the interruption problem (as signalled by the press of the yellow key) until a subject made a selection from the menu redisplayed following the interruption.

Efficiency was defined as the minimum number of frames required to find the answer to the question divided by the total number of frames accessed.

Subjects

Thirty subjects (6 female, 24 male) of average age 20.3 years (minimum 18, maximum 23) and with extensive experience of computers, were drawn from the undergraduate population of the Department of Electrical Engineering of Imperial College. Each was paid for their participation. Eighteen of the subjects had had experience with PRESTEL or other menu-based information systems.

Procedure

Subjects were asked to provide some demographic data and then to read a sheet explaining the experiment and the working of the system that they would be using. The Experimenter answered any questions and then stated: "There are two points I'd like to emphasise. Firstly, would you please answer the questions in order. Secondly, all questions can be answered using the data base: there are no 'trick' questions without answers !" The Experimenter then handed the subjects the question sheet and said "You may begin when you like". At the conclusion of the experiment the Experimenter asked the subjects to complete a questionnaire about the system that they had been using.

Experimental Design

The experiment employed a nested hierarchical random groups design, as illustrated in Figure 4. Of the 30 subjects, 15 were provided solely with the Standard Menu facility, while the other 15 were provided solely with the Selective Retreat facility. Each subject was required to find the answer to seven questions, three of which were interrupted as previously described and four which were not interrupted. The system maintained a complete time-stamped record of actions taken by subjects.

Data Analysis

The nested hierarchical design employed in the experiment required the use of a multi-variate analysis of variance model (MANOVA) that is capable of taking into account the jointly dependent variables that derive from questions that are Interrupted or non-Interrupted (the *INTERRUPT* factor) nested with questions that are Simple or Complex (the COMPLEXITY factor) nested with the independent conditions Selective Retreat or Standard Menu (the ENHANCEMENT factor). This design gives rise to eight (2x2x2) experimental units which are used in the analysis, as illustrated in Figure 4.

The MANOVA subprogram of the SPSS package (Statistical Package for the Social Sciences) allows for the analysis of data generated in a nested hierarchical design to give estimates of the effects of individual variables whilst taking into account the joint effects of the nesting of the variables, which is obviously crucial in this type of design. This subprogram was used to assess the data since the use of multi-variate analysis of variance allows the effect of several independent variables and their interactions to be considered using one test only. The problems associated with the use of multiple

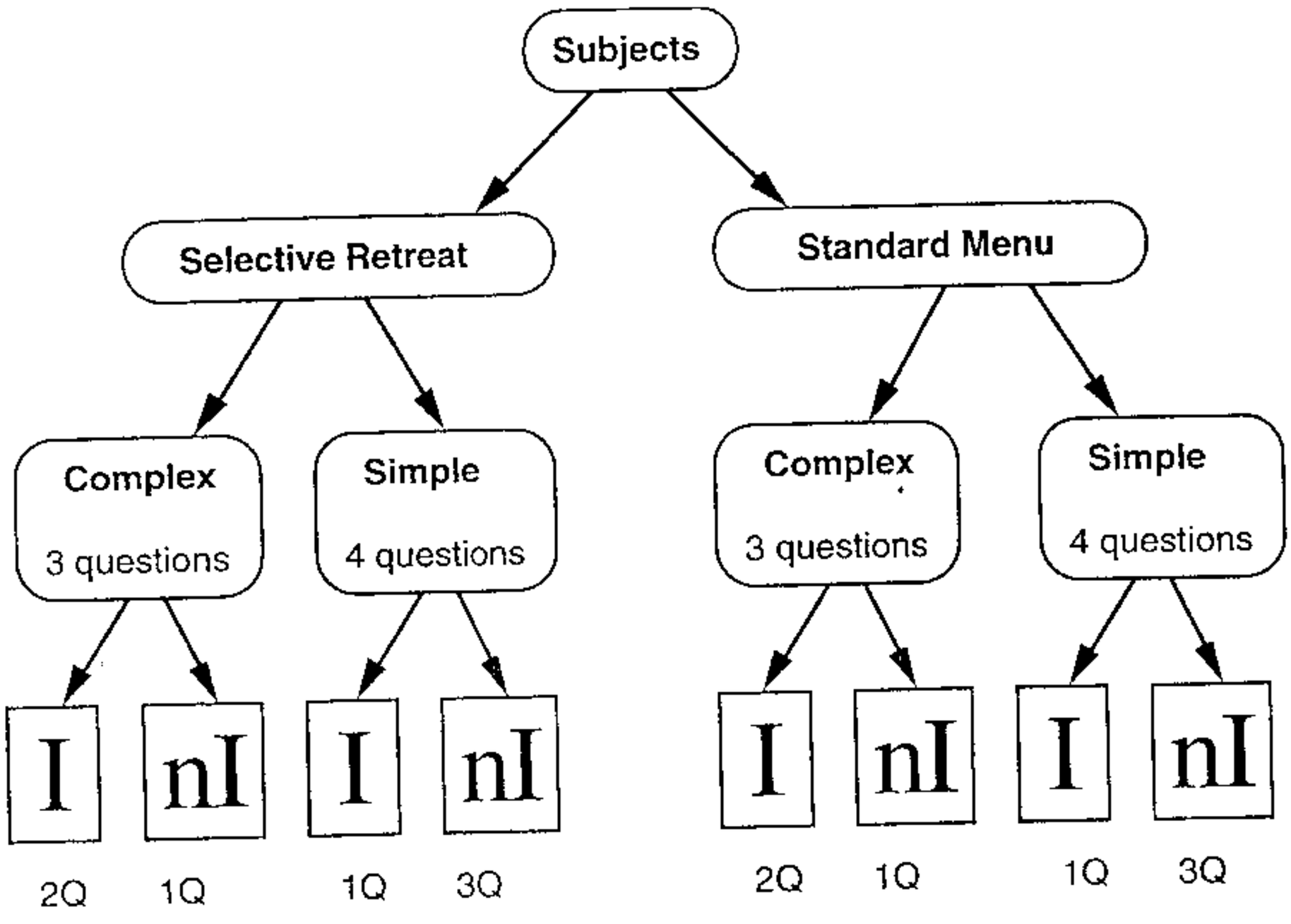

Figure 4 The nested hierarchical design used in the experiment
(**I** signifies interrupted questions, and **nI** non-interrupted questions)

t-testing, primarily the increased probability of making a Type I error, are therefore avoided.

Results

Each of the three main factors was able to be interpreted as a single entity using the MANOVA analysis, with the following results:

Interrupt

Interruptions were found to have a significant main effect (p=0.001) upon the ability of the subject to answer questions by reference to the data base.

Complexity

Question complexity was not a significant main effect (p=0.216)

Enhancement

There was no significant main effect for the enhancement factor (p=0.802)

Interactions

Interactions between the factors were also assessed using graphical and t-test analyses. These analyses indicated that there were no significant interactions between any combination of factors.

The above results are summarised in Table 1.

Interrupted questions

The mean number of frame accesses required to answer questions that were interrupted was 7.84 compared with the mean number of frame accesses required to answer questions that were non-interrupted of 6.26 (the same total number of

	Multivariate Analysis of Variance			
Factor	**F-value**	**df**	**p**	
INTERRUPT	3.6	(8, 404)	0.001	Main effect
COMPLEXITY	1.84	(4, 8)	0.216	No main effect
ENHANCEMENT	0.28	(2, 1)	0.802	No main effect

Table 1 Multivariate analysis of variance assessment of the effects of each of the factors within the experiment

	Interrupted Questions		non-Interrupted Questions		t	df	p
	mean	sd	mean	sd			
Frames Accessed	7.84	3.32	6.26	3.82	3.13	207	0.002
Search Time (seconds)	34.06	31.7	23.31	15.1	0.066	203	0.001

Table 2 Mean number of frames accessed, mean search times and t-test results for questions which were interrupted or non-interrupted

frames was required to answer each set of questions). The mean search time required to answer questions which were interrupted was 34.06 seconds and the mean search time required to answer questions that were non-interrupted was 23.31 seconds. Questions that were interrupted took significantly longer (p=0.001) and significantly more frame accesses (p=0.002) to solve than non-interrupted questions, as shown in Table 2. The average of these measures was increased respectively by 46% and 25% by the presence of interruptions.

Enhancement

The mean number of frames accessed by subjects in the Selective Retreat condition to answer all questions was 27.23 frames. Subjects in the Standard Menu condition used a mean of 28.03 frames to answer all questions. This difference was not significant. The mean search time for subjects in the Selective Retreat condition was 508 seconds and the mean search time for subjects in the Standard Menu condition was 456.5 seconds. Once again this difference is not significant. These data are presented in Table 3.

Post Interruption Activity

Overall, the first Post Interruption Pause (PIP) was a mean of 4.70 seconds (sd=2.9) and the fourth (last) PIP was a mean of 1.98 seconds. The difference was significant (t=4.51, df=58, p<0.001). However, a comparison of each PIP with the average time per selection in each of the interrupted questions shows that there was no significant difference between the mean screen scan period and the time taken to scan the screen

	Selective Retreat (SR)		Standard Menu (SM)		F	df	p
	mean	sd	mean	sd			
Frames Accessed	27.23	6.1	28.03	7.2	0.007	(1, 2)	0.943
Search Time (seconds)	508	162.3	456.5	154.1	0.066	(1, 2)	0.821

Table 3 Mean number of frames accessed and mean search times required to answer all questions and ANOVA results for subjects in the Selective Retreat and Standard Menu conditions

	PIP		Time per selection		t	df	p
	mean (seconds)	sd	mean (seconds)	sd			
Question 3	4.70	2.9	5.21	2.21	0.76	58	0.453
Question 4	3.31	3.29	4.62	2.7	1.68	58	0.098
Question 7	1.98	1.59	1.61	0.7	1.16	58	0.250

Table 4 Comparison of the Post-interruption Pause (PIP) and time spent viewing a screen for each of the interrupted questions

following an interruption. Thus, for the first PIP which had a mean of 4.70 seconds (sd=2.9), the mean time per selection (viewing a screen) in answering this question was 5.21 seconds (sd=2.21). Table 4 lists the PIPs and the comparison with the mean length of time spent viewing each screen with that particular question. The mean PIP for all interrupted questions was 4.01 seconds (sd=2.9) and the mean for non-interrupted questions was 1.98 seconds (sd=1.59). This difference was significant ($t=3.31$, $df=58$, $p<0.001$).

Efficiency

Efficiency was defined as the minimum number of frames required to answer the question divided by the total number of frames accessed. Interrupted questions were less efficiently answered. The mean efficiency for all interrupted questions was 0.62 (62%) and that for all non-interrupted questions was 0.78 (78%); this difference was significant ($p=0.002$).

In their separate categories, interrupted simple questions were less efficiently answered (efficiency = 57%) than non-interrupted simple questions (efficiency = 75%, $p=0.015$) and interrupted complex questions were less efficiently

	Interrupted Questions mean (seconds)	sd	non-Interrupted Questions mean (seconds)	sd	t	df	p
Efficiency Overall	0.62	0.23	0.78	0.3	3.28	58	0.002
Efficiency Simple	0.57	0.21	0.75	0.27	2.51	58	0.015
Efficiency Complex	0.71	0.22	1.61	0.3	2.58	58	0.012

Table 5 Mean efficiency for answering interrupted and non-interrupted simple and complex questions and questions overall

answered (efficiency= 71%) than non-interrupted complex questions (efficiency = 82%, p=0.012). Table 5 gives a summary of these data.

Discussion

The major conclusion of this study is that interruptions do have a significant effect upon the ability of subjects to answer questions by reference to the Carlton data base, both with respect to the number of frames accessed (an increase of 25%) and the search time required (an increase of 46%). This result is consistent with the conclusion of Gillie and Broadbent (1989) that an interruption has a significant effect if it must be attended to immediately. As pointed out earlier, the presentation period of the subject's selected frame was kept to a minimum (0.1 second) consistent with the provision of confirmatory feedback but negligible opportunity for rehearsal.

Our experiment addressed another issue raised by Gillie and Broadbent (1989) concerning the extent to which the availability of 'backtracking' (i.e., selective retreat) could influence the effects of interruption. In our experiment there was little difference to be seen in the use or otherwise of the Selective Retreat facility. This result is, in fact, not unexpected, since the questions posed did not require subjects to re-access the data base, a task that would have been supported by the Selective Retreat facility. Indeed, an earlier experiment (Field & Apperley, 1990) discovered that it was not until subjects were required to re-access the Carlton data base that the Selective Retreat facility gave rise to significant improvements in navigation. One could say that, in the experiment we report here, the subjects were given *questions to answer* rather than *problems requiring solution.*

Although the Post-Interruption Pause (PIP) time differed significantly from PIP1 to PIP4, comparisons of each PIP with the view time per frame for that particular interrupted question indicated that each PIP time was not significantly different from the average view time per frame for that particular question. This indicates that the decrease in PIP time was a reflection on increased facility with the system through practice or was a temporal position effect.

Anecdotally it was observed that the initial interruptions caused some surprise and that later interruptions were more expected. The fact that the PIP time was no different from the average screen scan times indicates that, in this experiment, the dependent variable PIP did not adequately reflect on interruption effects. However, increasing task or interruption complexity could yet prove that the PIP is an important dependent measure: it is certainly worth further investigation. This is also indicated in the Kreifeldt & McCarthy (1981) study.

Two other aspects of interruptions that would seem to be worthy of further study include increased interruption complexity through interrupting interruptions or requiring subjects to access other parts of the data base to deal with the interruption situation. Additionally, under these conditions, the beneficial effects of various navigational aids may also be demonstrated. In view of the lack of strong data concerning the effect of question complexity it would also be useful to address this issue in further research.

References

Gillie, T. and Broadbent, D. (1989) What makes interruptions disruptive ? A study of length, similarity and complexity, Psychological Research. 50, pp.243-250.

Kreifeldt, J.G. and McCarthy, M.E. (1981) Interruption as a test of the user-computer interface, in Proceedings of the 17th Annual Conference on Manual Control, JPL Publication 81-95, Jet Propulsion Laboratory, California Institute of Technology, pp.655-667.

Peterson, L.R. and Peterson, M.J. (1959) Short-term retention of individual verbal items, Journal of Experimental Psychology, 58, pp.193-198.

Field, G.E. (1987) Experimentus Interruptus, New York, ACM, SIGCHI Bulletin, 18, 4, pp.42-46.

Field, G.E. and Apperley, M.D. (1990) Context and Selective Retreat in Hierarchical Menu Structures, Behaviour and Information Technology, 9, 2, pp.133-146.

Heppe, D.L., Edmondson, W.H. and Spence, R. (1985) Helping both the novice and the advanced user in menu-driven information retrieval systems, in People and Computers: Designing the Interface, Cambridge, Cambridge University Press, pp.92-101.

Woolfe, R. (1980) Videotex - the new television-telephone information services, London, Heyden.

Biographical Note

Graeme Field is a psychologist, now woth Fletcher Holdings, Auckland. This paper is based on work carried out for his PhD study at Imperial College, London.

Robert Spence os Professor of Information Engineering at Imperial College, London. He has a long standing interest and experience in human-computer interaction research.