Psychol Res (1989) 50: 243-250

Psychological Research

What makes interruptions disruptive? A study of length, similarity, and complexity

Tony Gillie and Donald Broadbent

Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford, OX1 3UD, UK

Summary. Classic work on interruptions by Zeigarnik showed that tasks that were interrupted were more likely to be recalled after a delay than tasks that were not interrupted. Much of the literature on interruptions has been devoted to examining this effect, although more recently interruptions have been used to choose between competing designs for interfaces to complex devices. However, none of this work looks at what makes some interruptions disruptive and some not. This series of experiments uses a novel computer-based adventure-game methodology to investigate the effects of the length of the interruption, the similarity of the interruption to the main task, and the complexity of processing demanded by the interruption. It is concluded that subjects make use of some form of non-articulatory memory which is not affected by the length of the interruption. It is affected by processing similar material however, and by a complex mentalarithmetic task which makes large demands on working memory.

Introduction

The study of interruptions has a long history, going back to the classic experiments of Zeigarnik and Ovsiankina in the 1920s. In the 1940s interruptions were being implicated in pilot error and flying accidents (Fitts & Jones, 1947), with practical recommendations on how to reduce their assumed disruptive effect. More rencently, Kreifeldt and McCarthy (1981) and Field (1987) have used interruptions to compare different interfaces for calculators and computer databases. However, in both these more recent cases, the emphasis has been on comparing the merits of different designs, rather than on examining the effects of interruptions themselves. Current theories of memory and attention typically ignore the phenomenon of interruption; indeed, relatively little is known about the circumstances in which interruptions will prove disruptive. The experiments in the present paper therefore look specifically at the effects of different types and lengths of interruption on a computer-based task.

Zeigarnik, working with Lewin, carried out a series of experimental studies on the effects of interruptions on recall which were actually reported in this journal (Zeigarnik, 1927). Essentially, she showed that if people were interrupted during some tasks, but allowed to continue with others, then the interrupted tasks were recalled more often than the uninterrupted ones. The tasks used by Zeigarnik ranged from simple manual ones, such as stringing beads to more complex cognitive ones, such as solving puzzles. In a related series of experiments, Ovsiankina (1928) showed that if left to their own devices, subjects in such experiments would actually try to complete the interrupted tasks – a result that was taken to support Lewin's idea of a tension system needing to be discharged by the completing of a task. Since the original experiments there have been many attempts at replication and reinterpretation in line with a number of theories. (See: Van Bergen (1968) for a thorough review.) Although many experiments have been conducted to see whether the "Zeigarnik effect" is reliable (and arguably this is still not clear) none of them seems to have addressed the rather more interesting question of how easily people can resume what they were doing when the interruption finishes. Clearly, in practical terms, this is much more important than merely seeing whether people can remember which tasks were interrupted.

The recent studies by Kreifeldt and McCarthy (1981) and Field (1987) have used interruptions to compare different possible designs of interface with electronic devices. The strategy is simply to compare how easily people can pick up where they left off after an interruption when using these different interfaces. Kreifeldt and McCarthy used interruptions in this way to compare Reverse Polish Notation (RPN) and Algebraic Notation (AN) calculators. Two groups were each given problems of differing complexity to solve, and were interrupted after a set time (12 s) and asked to write down multiplication tables for 1 min. The performance of the two groups after interruption was then compared. Although there were differences in time to completion after an interruption between the groups using the RPN and AN calculators (with an advantage for the RPN group), there was also a general slowing down after an interruption in both groups. Unfortunately, the design did not allow control over the exact point in the task when an interruption occurred or over whether subjects started on the interrupting task when told to do so. Only four "key" times were taken for each subject, using a stopwatch, and the authors concede that they are unable to account for the effects of interruption in anything but the most general way.

Field (1987) employed a database traversal task, with subjects using a computer to answer specific easy or diffi-

Offprint requests to: D. Broadbent

cult questions. A variety of interruptions were used, such as completing a number sequence and looking up book titles, although it is not clear precisely which problems were interrupted by which tasks. Field claims to have shown a significant disruptive effect of the interruptions on users' post-interruption activity, although this is based solely on a comparison of the two experimental groups. The critical comparison of interrupted and non-interrupted problems is not reported. Conceivably, the difference reported is simply due to the experimental manipulation (how easily subjects could backtrack through previous choices) and not to the effect of the interruptions.

None of this previous work examines specifically why some interruptions are disruptive and some are not. The work of Kreifeldt and McCarthy suggests that similarity of the interruption to the main task may be a factor, although the relative length of the interruption may also be important – overall, mean time to completion for uninterrupted problems was only about 1¼ minutes while the interruption itself lasted 1 min! Field gives no examples of the number sequences subjects had to complete, though one could speculate that more complex sequences (e.g., a Fibonacci sequence) would be more disruptive than simpler ones. The present study will therefore examine length, similarity, and complexity as three possible explanations for the everyday observation that some interruptions are disruptive while others are not.

The task

The experiments are set in the context of a computer-based adventure game, where the subject needs to issue commands to the computer in order to achieve certain goals. Although unusual, this approach is not new. Lewis and Anderson (1985) used a similar approach in investigating problem solving. Essentially, the program models a small geographic area, with a number of locations that can be reached by issuing directional commands (e.g., north to location 1, southeast to location 2, and so on).

Design features common to all experiments

The subject is presented with a series of 12 separate problems, each of which has a list of items (either five or seven) that need to be "taken" (see Appendix). After seeing the original list, embedded in a plausible scenario, the subject moves in a specified direction. The computer then shows the present location (e.g., a baker), whether or not there are any objects present (e.g., bread), and lists other possible directions in which the subject can move, and where they lead. This list of options is presented in a different random order on each trial, to prevent the subject from simply visiting each location in turn. Pilot studies had shown that subjects (especially trained typists) were quite adept at remembering a position on the screen, and so their position in the list. As an additional precaution, each problem has an additional two "foil" locations, with items similar to those on the original list, so that subjects cannot simply visit all the locations shown to complete the task. In the experiments to be reported here, all locations had one, and only one, object present. The subject's task then, is to memorize a list of items, consider a list of locations, decide which is most likely to supply the next item required from the original list, move to that location, "take" the item, and then decide on the next location. When all the items on a list have been taken, the next problem is presented.

There were three distinct types of problem. First, "free order" problems, where the objects could be taken in any order. Second, "fixed logical" problems, where the objects had to be taken in a particular sequence and that sequence had some kind of logical ordering. Finally, "fixed arbitrary" problems, where the objects had to be taken in a particular sequence, but the order was completely arbitrary. The distinction between fixed- and free-order problems was discussed at some length by Miller, Galanter, and Pribram (1960), who refer to flexible and inflexible Plans. They speculated that inflexible Plans would use more working memory (their term) than flexible Plans, and would tend to be recalled more often after interruptions. Although their remarks were made in the context of the Zeigarnik effect, it was felt to be a sufficiently interesting point to test in the present experiments. Conceivably, fixed-order problems may differ from flexible ones in their susceptibility to interruptions. The reason for having two fixed-order problem types in this study was to allow a comparison of the effect of a fixed ordering per se with a fixed, but logical, sequencing. However, in none of the analyses reported here were there any differences between the three types of problem, nor did problem type interact with the effect of interruption. Accordingly, problem type will not be discussed further in this paper.

The presentation order of the problems, and the choice of which problems to interrupt is determined randomly for each subject, with the constraint that only three problems within each memory load are to be interrupted (one for each problem type). For five-item lists the interruption occurs after the subject has taken the third item; for seven-item lists it occurs after the fourth item. At the start of an interruption, the computer screen clears, and subjects are presented with the interrupting task. When the interruption is over, the screen is restored to its previous state.

Two different forms of help are available to the subject. First, they can ask to be "reminded" of the whole original list of items at any time. Second, they can ask for an "inventory", which lists the objects that have already been taken. If the subject forgets the commands available, the command "help" will redisplay them.

In order to minimize the amount of typing involved in the task, several keys are configured to issue whole commands (all the directional moves, reminders, and "take" are available as single keystrokes). Objects can be identified by the first two letters, although the program also accepts commands and object names typed in full. If a subject enters an incomplete command, the program prompts for the missing information. For example, the command "take" on its own would elicit the prompt "Take what?"

The program monitors the subject's progress through the experiment by recording which commands are given, where, and when. At the end of the experiment therefore, a full transcript of the subject's performance is available.

Experiment 1

Experiment 1 looks at the effect of 30 s of mental arithmetic as an interruption. Conceivably, any kind or length of interruption may disrupt this relatively demanding main task. Kreifeldt and McCarthy (1981) found that simply requiring subjects to write down multiplication tables had a detrimental effect on their ability to return to the main task of solving a problem with a calculator. One might predict that if retrieval of overlearned "arithmetic facts" (e.g., that $9 \times 9 = 81$) (see, e.g., McCloskey, Caramazza, & Basili, 1985) is enough to disrupt performance, then a task involving calculation should also prove disruptive. If it is not disruptive, then maybe similarity between the main and interrupting tasks, or the relative length of the interruption, is the major factor in determining which interruptions will be disruptive and which will not.

Method

Subjects. Ten subjects, four women and six men from the Oxford Subject Panel, all aged between 18 and 45, who were familiar with keyboards, were paid for their participation in this experiment.

Procedure. The experiment was run on an Acorn BBC 'B' computer with a Z80-second processor and 12-inch monochrome monitor. Subjects were tested individually. The instructions were presented by the computer, and subjects given the opportunity to ask questions. They were then given 2 min practice on simple mental arithmetic. The arithmetic problems were limited to two digit additions and subtractions, which were generated randomly by the computer. The problem was displayed on a single line (e.g., $27+56 = ?$). Subjects keyed their responses into the computer, and were told whether they were correct or not. After this, the subjects were given three practice problems for the main task, to familiarize them with the "adventure game" procedure. None of the practice problems was interrupted. Subjects then completed the main part of the experiment.

Results

The time spent on each problem was divided into two parts. For problems that were interrupted, this was simply the time before and after the interruption. For the remaining problems, the time at which an interruption would have occurred was calculated, and used to divide the total time accordingly. Before conducting an analysis of variance on these data, however, an adjustment was made to take account of the number of objects to be acquired in each part of the problem (for five-item lists, three in the first part and two in the second; for seven objects, four and three). The unit of analysis is therefore the time (in seconds) to take each object.

There is no main effect of interruption (mean not interrupted 18.51, interrupted 20.79), $F(1,9) = 2.66$, $p > .1$, and no effect of memory load (mean for five item lists 20.33, for seven item lists 18.98), $F(1,9) = 2.27$, $p > .1$. There is an effect of problem part however, with the time spent acquiring objects in the second part of any problem (mean 21.71) greater than that in the first (mean 17.57), $F(1,9) = 8.89$, $p < .02$.

In the absence of a main effect of interruption, the interaction of interruption with problem part would indicate a disruptive effect. However, in this experiment the interaction is not significant, $F(1,9) = 1.13$, $p > .3$.

An analysis of variance on the number of requests for help was conducted, by combining the number of requests for reminders with requests for an inventory. The results reveal the same pattern as the time data. There is no effect of interruption (mean not interrupted 0.46, interrupted 0.6), $F(1,9) = 2.50$, $p > .1$, and no effect of memory load (mean for five-item lists 0.48, for seven-item lists, 0.58), $F(1,9) = 2.06$, $p > .1$. There is a significant effect of the part of the problem, with more requests for help in the second part (mean 0.67) than in the first (mean 0.39), $F(1,9) = 14.39$, $p < .005$. As with the time data, the interaction between interruption and part of problem is not significant, $F(1,9) = 1.77$, $p > .2$.

In view of the uneven numbers of male and female subjects, an analysis comparing their performance was conducted. For all the experiments reported here, there is no main effect of sex, nor are any of the interactions with sex significant at the .05 level.

Discussion

Quite clearly, with this task, a short, simple, and dissimilar interruption does not have a disruptive effect. Some features of the main task might account for this. First, memory load was quite low at the point of interruption. Subjects had to remember only either two or three items from the original lists at the point of interruption. But it is interesting to consider the Kreifeldt and McCarthy (1981) result here. They found that a 1-min interruption of writing down multiplication tables did interfere with the task of solving a problem with a calculator. Arguably, all subjects would have needed to do in that experiment when interrupted was to remember their place in the calculation, rather than actually remembering the result so far (which the calculator does) or the steps completed. By contrast, in the present experiment subjects had a rather higher memory load, since they had to remember either specific items from the original list or the whole list and their position in it. Having a low memory load in the main taks then does not seem to guarantee immunity from the disruptive effects of an interruption, nor does having a higher memory load guarantee that there will be disruptive effects.

Second, the interruption was, by design, very short. Conceivably, with such a difference between the main task and the interruption, working memory (Baddeley 1986) can retain information from the main task while processing the interruption. Although it has been suggested that answers to single-digit mental arithmetic may occur spontaneously (Winkelman & Schmidt, 1974, but see also Ashcraft & Battaglia, 1978, and Groen & Parkman, 1972, for other accounts), mental arithmetic with more than two digits may well make use of the articulatory loop. In an unpublished study, we have shown that the error rate for mental arithmetic of this sort is significantly greater for a group with articulatory suppression than a comparable group performing a non-articulatory secondary task.

This suggests that subjects in this experiment are not relying exclusively on the articulatory loop to store information from the main task throughout the interruption, but rather on some other form of temporary storage. Even with traditional verbal learning tasks, it is now clear that a substantial amount of information is retained even when the articulatory loop and phonological store are disrupted (Gillie & Broadbent, submitted). It is not yet clear, however, what form this temporary storage takes.

The effect of problem part, with people slowing down in the second part of the problem is to be expected, given the design of this particular task. After taking an object, the subject returns to a list of locations, which is presented in a different random order on each occasion. At the start of a problem, when few objects have been taken, the chances of an univisited location being near the top of the list of options are greater than near the end of a problem, when most locations will already have visited. On average therefore, the time to scan the list and to decide on a location not yet visited will increase throughout the problem.

Experiment 2

Experiment 1 showed that an interruption with 30 s of simple mental arithmetic does not have a disruptive effect on the performance of the interrupted task. Further, it seems that subjects are relying on some form of temporary storage other than the articulatory loop to maintain their position in the main task. In order to see how long material in this temporary storage remains intact, the length of the interruption in this experiment is increased. Conceivably, something analogous to a simple trace-decay model of memory (e.g., Brown, 1958) will be adequate to explain why some interruptions are disruptive while others are not.

Method

Subjects. Ten subjects, three male and seven female, from the Oxford Subject Panel, aged between 18 and 45, who were familiar with keyboards, were paid for their participation in this experiment.

Procedure. The procedure remained the same as Experiment 1, with the same main and interrupting tasks. In Experiment 2, however, the length of the interruption was increased to 2.75 min.

Results

The time spent on each problem was divided into two parts, as described for Experiment 1 above. Analysis of variance shows that there is no main effect of interruption (mean not interrupted 19.26, interrupted 21.82), $F(1,9) = 3.34$, $p > .1$. There is a main effect of memory load with the high memory load (mean 22.92 s) taking longer than the low (mean 18.16 s), $F(2,18) = 7.09$, $p < .03$. There is no main effect of problem part, although numerically the effect is in the expected direction (mean for taking items in the first part 18.43, for the second part 22.65), $F(1,9) = 2.45$, $p > .1$. As in Experiment 1, the interaction of interruption and problem part is not significant, $F(1,9) = 3.07$, $p > .1$.

Analysis of variance on the help data shows the same pattern of results as the time data. There is no main effect for interruption (mean not interrupted 0.12, interrupted 0.15), $F(1,9) = 0.3$, $p > .5$. There is, however, a main effect of memory load, with more help requested at the higher memory load (mean 0.21) than the lower (mean 0.06), $F(1,9) = 5.44$, $p < .05$. There is no main effect of problem part (mean for first part 0.06, for second 0.21), $F(1,9) = 2.99$, $p > .1$, and the interaction of interruption and problem part was not significant, $F(1,9) = 0.04$, $p > .8$.

Discussion

The lack of a main effect of interruption, and the absence of the interaction of interruption with problem part shows that relatively long, but simple, interruptions do not disrupt the performance of the main task. So the length of an interruption on its own does not seem to be the critical factor in determining whether or not it will prove disruptive. Within the small range tested here, information stored in the non-articulatory temporary storage seems quite resistant to decay over time.

One curious feature of these results is the effect of memory load. Since the times used in the analysis have already been adjusted to allow for the number of items in each list, these data show that more time is needed for taking items from the longer lists. This seems particularly surprising given the small differences involved (two items). The explanation, however, may not be related to differences in retrieval strategies, but may be analogous to the problem-part effect noted above. That is, it takes longer on average to decide which location to visit near the end of a list than at the start. The same argument can be applied to explain why it takes longer to "take" items from the seven-item lists than from the five-item lists. The chances of an unvisited location being near the top of the list of options decreases the further through a list the subject gets. This effect of increasing decision time will necessarily be greater for longer lists. An alternative explanation may be in terms of the slightly increased chances of making an error with the longer lists (e.g., by going back to a previously visited location).

One possible reason for the lack of effect of interruption in Experiments 1 and 2 is that subjects have time to memorize the items in the main list in some way before starting to process the interruption. Although, when the computer screen cleared, subjects were immediately presented with a mental-arithmetic problem, they did not actually have to start processing the interruption until they were ready to do so, and this would have allowed time for some kind of rehearsal of the remaining items. This is unlike many interruptions in real life, where the interruption requires immediate attention. Experiment 3 therefore sets out to examine the effects of an interruption that demands immediate attention.

Experiment 3

If the lack of interruption effect in Experiments 1 and 2 is simply due to subjects being able to rehearse items from the main list before processing the interruption, then any task that prevents such rehearsal should show disruptive effects. If, on the other hand, the lack of effect is due to the interrupting task (mental arithmetic) being too dissimilar to the main task (a relatively complex memory task), then a memory-intensive, rather than a processing-intensive interruption could be expected to prove disruptive. Mental arthmetic is not suitable in either case, and so a different interrupting task is needed. Free recall was chosen, with subjects saying each word out loud as it was presented by the computer. Since there was no delay between the start of the interruption and the first word appearing, this should have prevented the subjects from rehearsing their position in the main task. If this task is disruptive as an interruption, then a further experiment can distinguish between the two possible explanations of similarity and lack of time for rehearsal.

Method

Subjects. Ten female subjects from the Oxford Subject Panel, aged between 18 and 45, who were familiar with keyboards, were paid for their participation in this experiment.

Materials. Eight lists of 32 words, with each list containing eight words from four categories were produced. To equate the lists for difficulty, the Battig and Montague (1969), category norms were used, and each list constructed so as to have a mean representativeness score of 3.98 (SD 0.07). (It must be said that the Battig and Montague norms are unlikely to be accurate now. There have been many cultural changes in the last 18 years, and the norms might well look very different if they were elicited from present-day British subjects.) The lists were randomized for presentation so that there were equal numbers of each category in each quarter of the list, and no two members of the same category adjacent.

Procedure. The problems used for the main task, and the procedure, were the same as those in Experiments 1 and 2. For the free-recall task, each word was displayed for 1.5 s, and there was a delay of 0.75 s between words. Subjects were required to say the words aloud as each appeared on the screen, and had 90 s for written recall after all the words had been presented. The length of this interruption was the same as in Experiment 2, that is, 2.75 min.

Subjects were given two practice lists for the free-recall task, as well as the usual practice on the main task. Which lists were used for practice, and the order of presentation of the lists, was randomized for each subject.

Results

The time spent on each problem was divided into two parts, as described for Experiment 1 above. Analysis of variance on the time to take each object shows a main effect of interruption, with performance slower on interrupted trials (21.71 s) than non-interrupted trials (18.35 s), $F(1,9) = 5.86, p < .04$. There is also a main effect of memory load, with subjects taking longer per object at the higher memory load (mean 21.75 s) than at the lower memory load (mean 18.31 s), $F(1,9) = 11.7, p < .01$. Finally, there is a main effect of part of problem, with performance in the second half (mean 22.43 s per object) being slower than that in the first (mean 17.63 s per object), $F(1,9) = 15.28$, $p < .004$. The interaction of interruption and problem part is not significant, $F(1,9) = 2.05, p > .1$.

Analysis of variance on the number of times subjects requested help shows that there is no main effect for interruption in the help data (not interrupted, mean 0.3, interrupted mean 0.46), $F(1,9) = 2.99, p > .1$. There is a main effect of memory load, $F(1,9) = 6.47, p < .04$, with more help being requested in the higher memory-load condition (mean 0.53) than the lower (mean 0.24). There is also a main effect of problem part, $F(1,9) = 18.44, p < .002$, with less help requested in the first half (mean 0.28) than in the second (mean 0.49). The interaction between memory load and interruption is significant, $F(1,9) = 5.26, p < .05$. Post-hoc analysis using the Newman-Keuls test shows that significantly more help is needed for the higher memory load when the problem is interrupted (mean 0.68), than either uninterrupted problems for the higher memory load (mean 0.38), or either type of problem for the lower memory load (mean interrupted 0.27, not interrupted 0.22), $p < .01$. No other differences between groups were significant.

One interesting feature of this experiment is that it presents the opportunity of examining whether people who perform differently on a traditional laboratory memory task – free recall – behave any differently on this rather more complex task. Subjects were divided into two groups using a median split, according to their performance on the first two practice lists (that is, where the lists were not acting as an interruption). When entered as a factor into the analysis of variance, there are no significant differences in level of disruption between those who scored high, and those who scored low, $F(1,8) = 0.34, p > .5$. It seems then that the disruptive effects noted here are quite general, and are not a simple reflection of individual differences in memory ability.

Discussion

Interruption with free recall has a disruptive effect on performance of the main task. This is clear in both the time and the help data. Of course, subjects may take longer either because they are trying to retrieve the original list or their position in it, or because they are spending time consulting the help system. (In practice, the help system was used very little, possibly as a consequence of the instructions, which stressed that help was only to be used "as a last resort.") In both cases, however, increasing times show that subjects are experiencing difficulties in retrieving items from the original list, and it seems fair to use this as a measure of disruption.

It is slightly puzzling that there is a main effect of interruption, and yet no significant interaction with problem part. This suggests that subjects are actually slower in interrupted problems before the interruption has occurred, as well as afterwards. Since the choice of problems to be interrupted and the presentation order of the problems are randomized for each subject separately, the explanation cannot be in terms of the operation of a systematic bias. It must be said, however, that although the interaction does not reach statistical significance, numerically the pattern is as expected. That is, the second part of interrupted problems takes longer than the first, and longer than either part of problems that are not interrupted. It is not clear, however, whether the disruptive effect noted here is due to the interruption preventing rehearsal, or to the similarity between the main and the interrupting tasks. Since there are many examples of interference amongst similar tasks in the standard memory literature (e.g., McGeoch & McDonald, 1931; Yntema & Mueser, 1962), a parsimonious explanation of the present results would be that there will be some disruption on return to a continuing task if an interruption occurs that involves processing similar material. This would certainly account for the Kreifeldt and McCarthy (1981) result, where both the main task and the interruption involved processing numeric information.

The issue of whether the opportunity for rehearsal can reduce or eliminate the disruptive effects of an interruption remains open, of course, even though the results from Experiment 1 suggest that people are not using the articulatory loop to retain information from the main task. The results from Experiment 1 may only show that interruptions that are similar to a main task will prove disruptive. Accordingly, a final experiment was devised in which the interruption would be dissimilar to the main task, more complex than the simple mental arithmetic used in Experiments 1 and 2, and yet would allow the subjects time to rehearse their position on the main task before processing the interruption.

Experiment 4

One of the conclusions from Experiment 2 was that the length of an interruption (within the small range tested) is not a crucial factor in determining whether or not it will be disruptive. Accordingly, for the final experiment, a short interruption was used. The task used was again mental arithmetic, but, unlike those in Experiments 1 and 2, the digits were coded as letters, and the subject had to "decode" the problems before performing the mental arithmetic. In common with Experiments 1 and 2 is the fact that subjects did not need to start processing the interruption until they were ready to do so.

Method

Subjects. Ten subjects, eight women and two men from the Oxford Subject Panel, all aged between 18 and 45, who were familiar with keyboards, were paid for their participation in this experiment.

Materials. The arithmetic problems were of the same kind as those used in Experiments 1 and 2, i.e., addition and subtraction of two-digit numbers, generated as required by the computer. In this task however, the numbers were coded as letters. To decode the problem, a displacement value was given (between 2 and 10) at the top of the screen, indicating which letter represents zero for that trial. For example, with a displacement value of 2, letter B = zero, C = one, D = two, and so on. With a displacement value of 3, letter C = zero, D = one, E = two, and so on. The original intention had been for subjects to visualize the alphabet to decode the problems, but pilot studies showed that this made the task unacceptably difficult. Accordingly, the alphabet was displayed in upper-case letters at the top of the screen throughout the problem. Subjects were instructed not to touch the screen at any time, and not to point to a particular letter to "mark" the zero point. Subjects keyed in the answer to the problem in digits, and were not expected to recode the answer into letters. Subjects spent a minimum of 30 s at the task, but were allowed to finish the problem they were working on before being returned to the main task. The mean time actually spent on interruptions was 51.08 s (SD 14.8 s).

Procedure. The problems used for the main task, and the procedure, were the same as those in the previous three experiments. Subjects were given a minimum of 5½ min practice on the coded mental-arithmetic task, and were allowed to finish the problem they were working on after this time had elapsed. They were again given practice on three problems from the main task.

Results

The time spent on each part of the problem was divided into two parts, as described in Experiment 1. An analysis of variance on the time to take each object shows no main effect of interruption (mean not interrupted 21.96 s, interrupted mean 22.60 s), $F(1,9) = 0.06$, $p > .8$. There is no main effect of memory load (mean low-memory load 20.92, mean high-memory load 23.64), $F(1,9) = 1.42$, $p > .2$, or of problem part (mean first part 20.34, mean second part 24.21), $F(1,9) = 2.17$, $p > .1$. The interaction between interruption and problem part is significant however, $F(1,9) = 6.04$, $p < .04$. Post-hoc analysis using the Newman-Keuls test shows that acquiring objects in the second part of problems that are interrupted (mean 26.02) takes significantly longer than in the first part of such problems (mean 19.18) and in both parts of uninterrupted problems (first part mean 21.50, second part mean 22.41).

Table 1. Interaction of problem part with interruption in Experiment 4

Problem part	Mean time for "taking" each object (seconds)	
	First	Second
Not interrupted	21.50	22.41
Interrupted	19.18	26.02

Finally, there is a significant interaction between memory load and interruption, $F(1,9) = 6.19$, $p < .04$. Post-hoc analysis using the Newman-Keuls test shows that problems at the higher memory load that are not interrupted take significantly longer (mean 24.94) than those at the lower memory load that are not interrupted (mean 18.98). No other means differ significantly.

An analysis of variance on the help data shows no significant effect of interruption (mean not interrupted 0.25, mean interrupted 0.3), $F(1,9) = 0.3$, $p > .5$. There is no effect of memory load (low memory-load mean 0.23, high memory-load mean 0.33), $F(1,9) = 1.36$, $p > .2$ or of problem part (first part 0.18, second part 0.38), $F(1,9) = 3.01$, $p > .1$. The interaction between interruption and problem part is not significant, $F(1,9) = 0.4$, $p > .5$. The interaction between memory load and problem part is significant, $F(2,18) = 4.56$, $p < .03$. Post-hoc analysis using the Newman-Keuls test, however, shows no significant differences between means.

Discussion

These results show quite clearly that having the opportunity to rehearse one's position in the main task does not automatically offer protection against the disruptive effect of an interruption. It is also interesting that a relatively short and dissimilar interruption has this disruptive effect. One possible interpretation might be in terms of complexity or of the amount of processing that the interruption requires. The evaluation of task complexity has, of course, been a vexed issue for a long time, with a large literature in the ergonomics field on mental workload, for example (Leplat, 1978). An interpretation in terms of complexity would be in line with work by Posner and Konick (1966), who essentially demonstrated that short-term memory for letters and digits is increasingly disrupted by more complex interpolated tasks. It would also fit with Broadbent's conception of an abstract working memory in the Maltese Cross model of memory (Broadbent, 1984).

That two such dissimilar tasks as recall of categorized lists (Experiment 3) and the coded arithmetic task (Experiment 4) should both show an effect of interruption suggests then that similarity of the material processed during the interruption to that in the main task is not the only factor in determining whether an interruption will prove to be disruptive. Complexity is clearly another factor that must be taken into account.

General discussion

Experiment 1, taken in conjunction with the result of Kreifeldt and McCarthy (1981) suggests that memory load at the time of an interruption is not a crucial factor in determining whether or not an interruption will be disruptive. In the Kreifeldt and McCarthy case, an apparently low memory load in the main task still led to disruption after the interruption. By contrast, in Experiment 1, subjects at both levels of memory load had a higher load than subjects in the Kreifeldt and McCarthy experiment, and yet were not disrupted by the interruption.

Experiment 2 demonstrates that within the range tested, the actual length of an interruption cannot be used to determine how disruptive it will be. This is in line with everyday experience, where some short interruptions can be so disruptive that the original task is forgotten, while some lengthy interruptions can be handled without causing any problems on return to the original task.

Experiment 3 shows that having to deal with an interruption that is similar to the main task and demands immediate attention is disruptive. It raises the question as to whether the opportunity to rehearse one's position in the main task guarantees immunity from the disrupting effects of an interruption. Experiment 4 demonstrates clearly that it does not, and further shows that similarity of the interruption to the main task is not crucial in determining whether or not an interruption will be disruptive. These results are summarized in Table 2.

One particular feature of this task is worth highlighting. After the subject takes each object, s/he returns to a list of location names that offer very strong cues to the items on the original list: e.g., Baker for bread, Butcher for meat, Post Office for stamps, and so on. This situation should be contrasted with that in which the state of the interrupted task makes it clear what the next step should be. It is clear that even with strong cues, after a disruptive interruption, subjects have difficulty in retrieving items from memory. It is interesting to contrast this with more traditional work on cuing recall (e.g., Tulving & Pearlstone, 1966), where cued recall is significantly better than non-cued recall. One major difference is that in the traditional experiments, the cue is typically presented at the same time as the item to be recalled. In these experiments the cues are not seen until after the list has been presented.

This series of experiments suggests then that the length of an interruption, and the opportunity to control the point at which the main task is stopped and the interruption started, are not important factors in determining whether or not an interruption will disrupt performance on return to the interrupted task. Rather, the nature of the interruption (in terms of similarity to the continuing task) and the complexity of the interruption (in terms of the amount of processing or memory storage required) seem to determine which interruptions will be disruptive and which will not.

Apendix

Five-item lists

You decide to go on a shopping expedition. You need some SHOES, a NEWSPAPER, some STAMPS, a VIDEO, and some BATTERIES (foils BOOK & RECORD)

You are planing a foreign holiday. You will need BROCHURES, TRAVELLERS CHEQUES, a PASSPORT, some SANDALS and a GUIDEBOOK (foils SUNTAN LOTION & SWIMWEAR)

You have a child who is going back to school. You will need to buy some SHIRTS, some SHOES, a GEOMETRY SET, a BRIEFCASE and some PAINTBRUSHES (foils BOOKS & SPORTS KIT)

You need to do some shopping. You need a VIDEO, a RECORD, a MAGAZINE, some STAMPS and HOLIDAY BROCHURES (foils BOOK & WRITING PAPER)

You are going to buy a house. You will need DETAILS OF HOUSES, a MORTGAGE, a LOAN, a LEGAL CONTRACT, and a REMOVAL VAN (foils INSURANCE POLICY & SURVEYORS REPORT)

You are going to buy a car and learn to drive. You need a DRIVING LICENCE APPLICATION, a CAR, some INSURANCE, DRIVING LESSONS, and a DRIVING TEST (foils BANK LOAN & MOT)

Seven-item lists

You need to start shopping for Christmas. You will need a TREE, some FAIRY LIGHTS, CARDS, STAMPS, WINE, a TURKEY and some VEGETABLES (foils CHRISTMAS CAKE & PAPER PLATES)

You are going to organise a party. You will need some FRUIT, WINE, CHEESE, INVITATIONS, POSTAGE STAMPS, PATE and some BREAD (foils PORK PIES & PAPER PLATES)

You are about to move into a new house, and need to buy some things for it. You need CURTAINS, PAINT, a BATHROOM SUITE, KITCHEN UTENSILS, FOOD, PLANTS, and LIGHTBULBS (foils FURNITURE & BEDDING)

You need to stock up for the weekend. You need some FRUIT, BREAD, a VIDEO, some FISH, WINE, CAKES and a new RECORD (foils CHEESE & STEAK)

You decide to build a house. You need a BANK LOAN, some LAND, PLANNING PERMISSION, a DESIGN FOR THE HOUSE, BRICKS, PAINT, and some FURNITURE (foils BATHROOM SUITE & LIGHT BULBS)

You have to arrange a wedding. You will need INVITATIONS, STAMPS, a CAKE, buy BRIDESMAIDS DRESSES, CHAMPAGNE, TAXIS, and to hire TOP HATS (foils RINGS & FLOWERS)

Table 2. Summary of experimental manipulations and their effect

Experiment	Duration	Similarity	Complexity	Opportunity to rehearse	Disruptive
1 (simple maths)	short	low	low	yes	no
2 (simple maths)	long	low	low	yes	no
3 (free recall)	long	high	?	no	yes
4 (coded maths)	short	low	high	yes	yes

Acknowledgements. Tony Gillie is funded, and Donald Broadbent employed, by the Medical Research Council U.K.

The authors are grateful to Dianne Berry and Stephen Payne for their helpful comments on an earlier draft of this article.

References

Ashcraft, M. H., & Battaglia, J. (1978). Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology: Human Learning and Memory, 4 (5),* 527–538.

Baddeley, A. (1986). *Working memory.* Oxford: Oxford University Press.

Battig, W. F., & Montague, W. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut Category Norms. *Journal of Experimental Psychology Monographs, 80,* Number 3.

Broadbent, D. E. (1984). The maltese cross: A new simplistic model for memory. *Behavioral and Brain Sciences, 7,* 55–94.

Brown, J. (1958). Some tests of the decay theory of immediate memory. *Quarterly Journal of Experimental Psychology, 10,* 12–21.

Fitts, P. M., & Jones, R. E. (1947). *Analysis of factors contributing to 460 "pilot-error" experiences in operating aircraft controls* (Memorandum Report TSEAA-694-12. Aero Medical Laboratory, Wright-Patterson Air Force Base, Dayton, Ohio).

Field, G. E. (1987). Experimentus Interruptus. *SIGCHI Bulletin, 19 (2),* 42–46.

Gillie, T., & Broadbent, D. E. (Submitted). Irrelevant speech, articulatory suppression, and the Stroop effect.

Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review,* 79, 329–343

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

Leplat, J. (1978). Factors determining work-load. *Ergonomics, 21 (3),* 143–149.

Lewis, M. W., & Anderson, J. R. (1985). Discrimination of operator schemata in problem solving: Learning from examples. *Cognitive Psychology, 17,* 26–65.

McCloskey, M., Caramazza, A., & Basili, A. (1985). Cognitive mechanisms in number processing and calculation: Evidence from dyscalculia. *Brain and Cognition, 4,* 171–196.

McGeoch, J. A., & MacDonald, W. T. (1931). Meaningful relation and retroactive inhibition. *American Journal of Psychology, 43,* 579–588.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior.* London: Holt, Rinehart, and Winston.

Ovsiankina, M. (1928). Die Wiederaufnahme unterbrochener Handlungen. *Psychologische Forschung, 11,* 302–379.

Posner, M. I., & Konick, A. F. (1966). On the role of interference in short-term retention. *Journal of Experimental Psychology, 72 (2),* 221–231.

Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior, 5,* 381–391.

Van Bergen, A. (1968). *Task interruption.* Amsterdam: North-Holland.

Winkelman, J. H., & Schmidt, J. (1974). Associative confusions in mental arithmetic. *Journal of Experimental Psychology, 102,* 734–736.

Yntema, D. B., & Mueser, G. E. (1962). Keeping track of variables that have few or many states. *Journal of Experimental Psychology, 63 (4),* 391–395.

Zeigarnik, B. (1927) Das Behalten erledigter und unerledigter Handlungen. *Psychologische Forschung, 9,* 1–85.

Received December 5, 1988.