Br. J. Psychol. (1975), **66**, 3, *pp.* 325–331
Printed in Great Britain

CONTEXT-DEPENDENT MEMORY IN TWO NATURAL ENVIRONMENTS: ON LAND AND UNDERWATER

By D. R. GODDEN and A. D. BADDELEY
Department of Psychology, University of Stirling

In a free recall experiment, divers learnt lists of words in two natural environments: on dry land and underwater, and recalled the words in either the environment of original learning, or in the alternative environment. Lists learnt underwater were best recalled underwater, and vice versa. A subsequent experiment shows that the disruption of moving from one environment to the other was unlikely to be responsible for context-dependent memory.

The philosopher John Locke cites the case of a young man who learned to dance in a room containing an old trunk. Unfortunately, however, 'the idea of this remarkable piece of household stuff had so mixed itself with the turns and steps of all his dances, that though in that chamber he could dance excellently well, yet it was only while that trunk was there; nor could he perform well in any other place unless that or some other trunk had its due place in the room' (Locke, cited in Dennis, 1948, p. 68).

The belief that what is learnt in a given environment is best recalled in that environment has of course been a useful standby for detective story writers from Wilkie Collins onwards, although the empirical evidence for such a belief is somewhat equivocal. Farnsworth (1934) and Pessin (1932) were both unable to obtain a context-dependent memory effect. A later study by Jensen *et al.* (1971) was more successful, but a recent unpublished study by Hitch (personal communication) failed to observe any effect. An alternative approach to the context-dependent phenomenon utilizes a retroactive interference (RI) design in which material learned in one environment is followed by a second set of material presented in either the same or a different environment, which in turn is followed by a recall test on the original material. This final test itself may be in the initial environment or in the interpolated environment. Using this design, Bilodeau & Schlosberg (1951) found that an interpolated list caused only half as much RI when it was learned in a room differing markedly from that in which original learning took place. Comparable results were obtained by Greenspoon & Ranyard (1957), and by Zentall (1970). However, Strand (1970) has presented evidence suggesting that the inferior retention observed when the recall environment is different results not from the different context *per se*, but from the disruption that occurs when the subject moves from one environment to the other. She required the subjects in her control conditions, in which learning and recall were in the same environment, to leave the room and have a drink of water from a drinking fountain before beginning the recall phase of the experiment. Under these conditions she found no reliable difference between subjects who learned the interfering material in the same and those who learned in a different environment.

The evidence for context-dependent memory is therefore far from convincing. Furthermore, a number of the studies which have obtained effects have used extremely artificial environments, suggesting that any effect observed may not generalize beyond the conditions under which the experiment was run. Dallett & Wilcox (1968), in one of their conditions, required their subjects to stand with their heads inside an oddly shaped box containing flashing lights of different colours (two subjects had to be excused due to nausea), while in a further study Rand & Wapner (1967) strapped their subjects to a board which was then rotated so as to keep the subject either supine or erect. As Brunswik (1956) has pointed out, it is important for the psychologist to check the 'ecological validity' of his results by ensuring that they are applicable to the real world, and are not limited to the artificial conditions of the psychological laboratory. The experiments to be described are therefore concerned with investigating the phenomenon of context-dependent memory in two natural environments. Egstrom *et al.* (1972) observed that divers had considerable difficulty in recalling material learnt under water, and since recognition memory was not impaired in this way, they suggested that the defect was probably due to a context-dependent memory effect rather than to differential learning underwater. Unfortunately, since the appropriate controls were not included, such a conclusion was purely speculative. However, the underwater environment does present a particularly good example of a natural environment which differs dramatically from that on the surface. The diver underwater is weightless, has restricted vision, and subjectively is in an environment which is as different from the surface as any he is ever likely to experience. Divers were therefore asked to learn word lists both on land and underwater and subsequently recalled either on land (Dry) or underwater (Wet). Each diver performed under all four possible conditions: DD (Learn Dry, Recall Dry); DW (Learn Dry, Recall Wet); WW and WD. Should the phenomenon of context-dependent memory exist under these conditions, performance when learning and recall took place in the same environment (DD and WW) should be significantly better than when recall took place in a different environment to that of learning (DW and WD).

EXPERIMENT I

Method

Subjects

Eighteen subjects were tested, comprising 13 male and five female members of a university diving club.

Apparatus

Five lists of words, each consisting of 36 unrelated, different, two- and three-syllable words chosen at random from the Toronto word bank, were constructed and subsequently recorded on tape (see *Procedure*).

Two Diver Underwater Communication (DUC) sets were used. A DUC set consists of a surface-to-diver telephone cable, terminating in a bone transducer, which, placed on the diver's mastoid, enables both surface-to-diver and diver-to-surface communication. The two DUC sets were slightly amended such that taped material, monitored by the surface operator, could be presented directly to the subject using a cassette tape-recorder. Twin transducers on each set allowed two separate pairs of subjects to be tested during the same period. Weighted formica boards, sealed with fablon, enabled subjects to record responses in pencil both above and underwater. Subjects used standard SCUBA breathing apparatus and diving equipment of various designs dictated by personal preference.

Procedure

All instructions and stimuli for each experimental session were recorded on tape. Efficient auditory perception of stimuli by a submerged diver using SCUBA apparatus is seriously impaired by the noise of his breathing. The presentation of the material was therefore grouped, so as to allow the diver to adopt a comfortable breathing rate which did not interfere with his auditory perception. Thus, each list was presented in blocks of three words. Within each block, the words were spaced at 2 sec. intervals. Between each block, a 4 sec. interval enabled subjects to exhale, inhale and hold their breath in readiness for the presentation of the next block, and so on.

Each tape began with an explanation of this breathing procedure, followed by a 'breathing pattern' section, to ensure that subjects were breathing correctly and in rhythm before the first word of the list appeared. This section consisted of nine spoken presentations of the letter *z*, in three blocks of three, and with identical spacings to those of the words in the list itself. Immediately after each block of *z*'s, subjects heard the command 'breathe'. The presentation of the word list followed on naturally in rhythm, and the command to breathe was then dropped.

On each tape (one for each of the 16 condition/list combinations), the relevant list was presented twice. Between the first and second presentations, a gap of 10 sec. allowed subjects a short rest with unconstrained breathing. The second presentation was again preceded by the breathing pattern section.

To eliminate possible primary memory effects, the second presentation of each list was followed by 15 digits which subjects were required to copy at a rate of two seconds per digit. This was followed by the next instruction (e.g. 'Ascend to the shore station'), and a 4 min. delay. This recurred in all conditions and was necessary to enable subjects to comply safely with the relevant instruction. They were then instructed to write down in any order as many of the words from the list as they could remember. Two minutes were allowed for this, after which the session was terminated.

The original 16 subjects were split at random into four groups of four. Prior to the first experimental session, all subjects underwent a practice session, comfortably seated around a table. During this, they first practised the breathing technique, then the task itself, using a practice list.

Pairings of the remaining four lists $L_1 \ldots L_4$ with the four conditions, and the temporal orderings of the conditions for each of the four groups, were arranged according to a Graeco-Latin square design. Subjects experienced one condition per diving session, and the sessions were separated by approximately 24 hr. The design was such that each group experienced conditions and lists in different orders, that a given condition/list pair was never administered to more than one group, and that lists and conditions had equal representation on each experimental session.

Testing took place at open water sites near Oban, Scotland, in most cases as soon as subjects returned from their scheduled dive for the day. The latter consideration helped to ensure that subjects began each session in roughly the same state, that is, wet and cold. They were tested in pairs which remained the same across all sessions.

Subjects in environment D (Dry) sat by the edge of the water, masks tipped back, breathing tubes removed, and receivers in place. In environment W (Wet), subjects dived to approximately 20 ft., taking with them their formica board and two pencils, and with their receivers in position. Heavy weighting enabled them to sit on the bottom, and the session began after a verbal signal to the surface operator signified their readiness.

Due to a series of technical difficulties, two subjects were eventually dropped from the programme. They were subsequently replaced by another pair who were tested at a freshwater site.

Results

Mean recall scores and standard deviations are shown in Table 1. An analysis of variance showed there to be no significant main effect on recall performance due to either the environment of learning or that of recall. The interaction between

Table 1. *Mean number of words recalled in Expt. I as a function of learning and recall environment*

	Recall environment				
	Dry		Wet		
Learning environment	Mean recall score	S.D.	Mean recall score	S.D.	Total
Dry	13·5	5·8	8·6	(3·0)	22·1
Wet	8·4	3·3	11·4	(5·0)	19·8
Total	21·9	—	20·0	—	—

the effect of learning and recall environment was, however, highly significant ($F = 22{\cdot}0$; d.f. = 1, 12; $P = < 0{\cdot}001$). Thus the effect on recall of the environment of recall depended on the environment of original learning. No other interactions proved significant. Finally, to look in more specific detail at the data, a Wilcoxon matched-pairs, signed-ranks test showed that, for learning in environment D, recall in environment D was better than recall in environment W ($P < 0{\cdot}005$) while for learning in environment W, recall in W was better than recall in D ($P < 0{\cdot}025$). There was no significant difference in recall between conditions DD and WW, nor between conditions DW and WD.

Discussion

The results of Expt. I are clearly in line with the context-dependent memory hypothesis: what was learned under water was best recalled under water and vice versa. Before accepting a context-dependent interpretation, however, some possible shortcomings of the experiment should be considered. The divers were on a pleasure-diving holiday at the time of the experiment, and it is entirely due to their good will in agreeing to participate, and tolerance in accepting the subsequent demands of the experiment that the latter was completed at all. Nevertheless, since the divers were in no way committed to the experiment, some limitations on what could reasonably be done were experienced. Thus, there was no control over the time of day at which subjects were tested. In addition the experiment had to be run each day at the diving site chosen by the subjects, rather than at a constant location. Diving expeditions are notoriously difficult to organize and run smoothly; in an equipment-intensive operation which depends strongly on such local conditions as weather, fitness, etc., something will usually disrupt planned routine. Some divers may not, for medical reasons, dive each day. A dive may have to be aborted due to equipment failure, and so on. These problems were experienced to some degree, as was the unexpected. One diver was nearly run over during an underwater experimental session by an ex-army, amphibious DUKW. Thus it proved impossible to complete the session in four successive days, and the time between sessions varied both within and between subjects. None the less, under realistic open water conditions, and even subject to the above problems, a highly significant interaction between the environment of learning and that of recall emerged.

Before concluding that this result reflects the effect of context-dependent memory three alternative explanations should be considered. The first of these assumes differential cheating between the four conditions, with subjects in the WW group copying items down before the recall signal. Unfortunately, since it was not practicable to combine monitoring of both the subject and the shore station, this is a logical possibility. It is, however, unlikely for the following reasons; (1) had this been the case, higher level of WW performance might have been expected unless subjects were cheating in a particularly restrained manner; (2) it would be difficult to cheat without the awareness and connivance of one's diving partner; and (3) this particular group of subjects had virtually all had previous involvement in an underwater research project of their own, several were in fact research scientists, and social pressures against cheating would be considerable. Nevertheless, differential, restrained collaborative cheating remains a possible explanation. The second alternative explanation assumes that subjects rehearse during the unfilled interval between presentation and test, the degree of rehearsal being greater in the WW and DD conditions than in WD and DW, where subjects might be expected to be distracted by the procedure of getting into or out of the water. Thirdly, there has been a suggestion (Strand, 1970) that context-dependent effects may not be due to environmental change *per se*, but instead to disruption caused by taking the subject from one environment to the other in the context-change conditions. Thus further discussion of the apparent context-dependent effect found here will be postponed until after the report of the next experiment.

Experiment II

This tests both the differential rehearsal and disruption hypotheses by comparing the standard DD condition used in Expt. I with a modified DD condition in which subjects are required to enter the water and dive during the 4 min. delay between presentation and the subsequent dry test. Both the differential rehearsal and disruption hypotheses should predict poorer retention in this condition, whereas a context-dependent hypothesis would not predict a decrement, since both learning and recall occur in the same environment.

Method

Subjects

Sixteen members of the Scottish Sub-Aqua Club served as voluntary subjects.

Apparatus

Subjects when in the water breathed using a snorkel tube, since this was administratively more convenient than SCUBA equipment. For the limited diving involved the two types of equipment may be regarded as equivalent. All other apparatus including word lists was the same as for Expt. I.

Procedure

As before, stimuli and instructions were all presented via tape. In condition *n* (non-disrupted), tapes employing the relevant lists under condition DD for Expt. I were used. For condition *d* (disrupted), the disruption instruction (to enter the water, swim a short distance, dive to approximately 20 ft., and return to the original position) was dubbed on to copies of the above tapes.

Subjects were assigned randomly to four groups as they volunteered. Prior to the experimental session, subjects practised the task and were then tested at an open water diving site near Oban, Scotland, immediately on their return from their scheduled day's dive.

A within-subject design was adopted. Groups 1 and 2 experienced condition *n* on the first session while groups 3 and 4 were presented first with condition *d*. List *A* occurred on the first session and List *B* on the second, for groups 1 and 3; List *B* was used on the first session and List *A* on the second, for groups 2 and 4. The two sessions were separated by approximately 24 hr.

Insufficient volunteers were found at Oban to complete the design. Consequently, a further three subjects from Stirling University were tested on a subsequent occasion at a freshwater site.

Results

In condition *n*, the mean number of words recalled was 8·44 ($\sigma = 4{\cdot}1$), while condition *d* produced a mean recall score of 8·69 ($\sigma = 3{\cdot}2$). The marginal difference, in the opposite direction to that predicted by the disruption hypothesis, did not approach significance.

Discussion

Scores obtained in Expt. II were lower than those obtained during the comparable DD conditions of Expt. I. Two possible reasons for this arise. Considerable background noise from other nearby diving groups persisted at a relatively constant rate throughout the experiment; this was not generally the case in Expt. I. Furthermore, most of the volunteers, unlike Expt. I subjects (largely medical students), were unfamiliar with testing situations and experimental procedures. Whatever the reason, the criticism might arise that possible effects of disruption might be masked by a floor effect. However, the two mean scores, both between 8 and 9, represent about 25 per cent recall, a reasonable level at which one might expect an effect to be measurable if it exists. Secondly, if 'high scorers' are identified as those subjects who scored higher than the mean under either condition, it is found that this sub-group recalled a mean of 10·1 words in condition *n* and 10·4 in condition *d*, suggesting that the absence of a disruption effect characterizes both high and low scores. Thus it can be concluded that under open water conditions similar to those for Expt. I, disruption was not a significant factor.

Conclusions

The effect of the environment or context of recall on performance depends upon the environment of learning. Recall is better if the environment of original learning is reinstated. This is unlikely to be due to disruption, and one can be reasonably confident that it is a truly context-dependent phenomenon. The results were obtained under real conditions, away from the laboratory, and using natural environments. The open water setting for the project limited the design such that two separate experiments were needed. However, within the limitations thus imposed, it appears that the phenomenon of context-dependent memory not only exists, but is robust enough to affect normal behaviour and performance away from the laboratory.

We acknowledge the invaluable help rendered by the Clydebank, East Kilbride, and Stirling University Sub-Aqua Clubs. In particular we would like to thank the United London Hospitals Diving Group. The experiment was supported by a grant from the Medical Research Council.

References

BILODEAU, I. M. & SCHLOSBERG, H. (1951). Similarity in stimulating conditions as a variable in retroactive inhibition. *J. exp. Psychol.* **41**, 119–204.

BRUNSWIK, E. (1956). *Perception and the Representative Design of Psychological Experiments.* Berkeley: University of California Press.

DALLETT, K. & WILCOX, S. G. (1968). Contextual stimuli and proactive inhibition. *J. exp. Psychol.* **78**, 475–480.

DENNIS, W. (ed.) (1948). *Readings in the History of Psychology.* New York: Appleton-Century-Crofts.

EGSTROM, G. H., WELTMAN, G., BADDELEY, A. D., CUCCARO, W. J. & WILLIS, M. A. (1972). Underwater work performance and work tolerance. Report no. 51, Bio-Technology Laboratory, University of California, Los Angeles.

FARNSWORTH, P. R. (1934). Examinations in familiar and unfamiliar surroundings. *J. soc. Psychol.* **5**, 128–129.

GREENSPOON, J. & RANYARD, R. (1957). Stimulus conditions and retroactive inhibition. *J. exp. Psychol.* **53**, 55–59.

JENSEN, L. C., HARRIS, K. & ANDERSON, D. C. (1971). Retention following a change in ambient contextual stimuli for six age groups. *Dev. Psychol.* **4**, 394–399.

PESSIN, J. (1932). The effect of similar and dissimilar conditions upon learning and relearning. *J. exp. Psychol.* **15**, 427–435.

RAND, G. & WAPNER, S. (1967). Postural status as a factor in memory. *J. verb. Learn. verb. Behav.* **6**, 268–271.

STRAND, B. Z. (1970). Change of context and retroactive inhibition. *J. verb. Learn. verb. Behav.* **9**, 202–206.

ZENTALL, T. R. (1970). Effects of context change on forgetting in rats. *J. exp. Psychol.* **86**, 440–448.

(*Manuscript received* 22 *March* 1974; *revised manuscript received* 5 *August* 1974)