Interevent differences in event memory: Why are some events more recallable than others?

RONALD L. COHEN, MICHELE PETERSON, and TONI MANTINI-ATKINSON
Glendon College, York University, Toronto, Canada

Interitem differences in the free recall of action events were studied in five experiments. The action events were presented in three different formats: minitasks performed by the subjects in response to verbal instructions from the experimenter (SPTs), minitasks performed by the experimenter (EPTs), and task instructions (TIs). Not only were reliable interevent differences in recall probability demonstrated within each format, but these differences tended to correlate across formats, especially between the SPTs and EPTs; thus, a highly recallable SPT also tended to be a highly recallable EPT. Attempts to explain interitem recall differences in terms of differences in familiarity, vividness, and the availability of environmental cues were largely unsuccessful. An experimental analysis of the action events into action and object components showed the recall probabilities of our events to be mainly dependent on the recall probabilities of their action components, with only a minor dependence on the recall probabilities of their object components.

One question asked of psychologists by laypersons is Why is it easy to remember some things but not others? Cognitive psychologists can answer this question by referring to the literature on the free recall of word lists, where interitem differences are explained in terms of encoding differences. This explanation is based, at least in part, on the demonstrated sensitivity of word memory to encoding variables. For example, words that are semantically encoded are generally better recalled than words that are phonemically or graphemically encoded (Craik & Tulving, 1975); words that are visually imaged are more likely to be recalled than words that are not (Paivio, Yuille, & Rogers, 1969); and so on. Consequently, the question of interitem differences in word recall can be answered by reference to the encoding properties of the items. Words that are readily imaged (e.g., concrete words) or that can be readily encoded in terms of existing cognitive structures (e.g., familiar words) are generally better recalled than items that do not possess these encoding advantages (viz., abstract words and unfamiliar words, respectively).

There are, however, many memory events that are not words, and an obvious question for cognitive psychologists is whether our knowledge of word-list memory can be generalized to enable us to explain interitem differences in recall for nonword events. If the answer to this question proves to be negative, then the question that opened this introduction (i.e., Why is it easy to remember some things but not others?) would have to be readdressed. The aim of this paper is to examine memory for some types of nonword events, specifically for action events, within the context of these two questions. In the

first part of the paper, we explore the possibility of using a word-list memory approach to explain interitem differences in the recall of action events. In the second part, we consider two explanations for interitem differences in action-event recall, explanations that do not stem from our knowledge of word-list memory.

Memory for action events has been studied in our laboratory using three basic procedures. In one procedure the experimenter gives the subject a series of minitask instructions, such as "clap your hands" or "break the toothpick." Following each instruction, the subject performs the appropriate task. In this format the events have been designated subject-performed tasks, or SPTs (Cohen, 1981). The second procedure is similar to the first, with the modification that each instruction is followed by the experimenter's performing the task while the subject looks on (experimenter-performed tasks, or EPTs). In the third procedure the task instructions are presented without any overt enactment (task instructions, or TIs).

Memory for SPTs has been found to differ from that for words in that the recall of SPTs appears to be minimally affected by standard laboratory encoding manipulations. For example, the levels of processing manipulation (Craik & Tulving, 1975), which has such a reliable effect on word recall, failed to produce an effect on SPT recall (Cohen, 1981). Designating some words as being important to recall, during acquisition, markedly increased their likelihood of being recalled, as compared with words not given an importance designation. This manipulation had only a minimal effect on SPT recall (Cohen, 1983). Also, reducing the rate of presentation, which has a considerable effect on word recall, presumably because of the increased time available for acquisition, was found to have no effect on SPT recall (Cohen, 1985). The general conclusion drawn from these earlier studies is that SPT memory, unlike word memory,

is nonstrategic in the sense that recall level has little or no dependence on acquisition strategies. This conclusion also received support from studies that have looked at population differences in SPT recall. Retarded individuals, children, and the aged are generally supposed to perform poorly on strategic tasks. Yet intersubject differences in SPT recall appear to be unrelated to intellectual capacity (Cohen & Bean, 1983) or to age (Bäckman & Nilsson, 1984, 1985; Cohen & Stewart, 1982).

Bäckman and Nilsson (1985) have suggested that SPTs have an advantage over verbal events because they possess an inherent richness. SPTs are, for example, multimodal, whereas verbal events are unimodal. Verbal events may be enriched through recoding, of course, which would bring them more in line with SPTs. This view of action events and verbal events is consistent with that put forward by Cohen and his colleagues, although it is somewhat more explanatory. Thus SPTs may exhibit nonstrategic properties, because these events produce such a rich memory trace without recoding. Memory for TIs may be considered strategic, because the richness of the traces in this case depends on the quantity and quality of recoding applied to these events. By Bäckman and Nilsson's criteria, EPTs must also be regarded as being inherently rich, although the subject has more control over the amount of attention paid to these events than in the case of tasks that have to be performed (SPTs). Given these apparent differences between word memory and SPT memory, it would not be too surprising if current explanations for interitem differences in word recall did not prove adequate to explain interitem differences in action-event recall.

Cohen (1983) has suggested that memory events may be ordered on a continuum of sensitivity to encoding variables, with words at the high-sensitivity end of the continuum and SPTs at the low end. Although the other two formats of action events have not been studied as systematically as the SPTs, the EPTs appear to lie closer to the SPTs on the continuum than do the TIs. All three variants of action events were included in the present investigation, partly so that we could study interitem differences in the three cases, but also so that we could further explore the relationship among the three event formats.

As a first step in our studies, we established recall probabilities for a pool of action events in Experiments 1 and 2. Subsequent to these experiments, action events having the highest and lowest recall probabilities were selected for further study.

Three main approaches to the pattern of interitem differences were explored, only one of which was derived from the word memory literature. This latter approach was used in Experiment 3, in which correlations were sought between event recall probability and factors that previously have been demonstrated to correlate with word recall probability, namely familiarity, vividness, and, in the case of the TIs, imageability.

The second approach is based on the possible influence of the physical environment on recall. In a free recall test of words, each word, following its presentation, is completely removed from the environment. In tests of action-event memory, the events leave traces not only in the subjects' memory, but also in the physical environment. For example, the action event "point to the door" is not completely removed following presentation, since the door remains. These event remnants could then serve as cues during retrieval. Pursuing this argument provides a basis for classifying the action events into two types, namely those that require the use of an extra-environmental object (object events), such as a toothpick for "break the toothpick," and those that do not (no-object events), such as "point to the door." Since the extra-environmental objects are removed from the subject's view immediately following event presentation, the environment should provide poorer retrieval cues for object than for no-object events. This conceptual difference between object and no-object events was used as a basis for testing the possibility that interevent differences in recall have some dependence on the availability of environmental cues at retrieval. This possibility was examined in Experiments 1 and 2, and also in Experiment 4.

An additional characteristic of our action events, not shared by words, is that each event consists essentially of two potentially separate components, an action component and an object component. Furthermore, the demonstrated importance of the motor aspect of action events both in SPT memory (Bäckman, 1985) and in memory for action sentences (Engelkamp & Zimmer, 1984; Saltz & Donnenwerth-Nolan, 1981; Zimmer & Engelkamp, 1985) suggests that an experimental separation of the object and action components might be worth pursuing in the interevent differences context. Our third approach, then, involved an experimental separation of the action and object components of the action events. This was done in Experiment 5 in order to determine whether recall probabilities of action events are determined by the nature of the actions, or of the objects, or of the unique combination of the two.

EXPERIMENT 1

In Experiment 1, we established recall probabilities by having groups of subjects recall lists of action events, and recording the proportion of subjects recalling each event. The reliability of these recall probabilities was measured by repeating the procedure with other groups of subjects and calculating test–retest reliabilities. In addition, correlations were calculated between the recall probabilities of the action events presented in the three different formats. Since EPTs are assumed to lie closer to SPTs than do TIs on Cohen's (1983) strategic continuum, we expected that correlations involving recall probabilities would be higher between SPTs and EPTs than between SPTs and TIs.

Method

Subjects. Subjects were 108 volunteers, ranging in age from 18 to 50 years, who were visitors to the Ontario Science Centre in Toronto.

Materials. A sample of 48 action events was selected from an expanded version of Cohen's (1981) pool. Some of the events involved the manipulation of an object (e.g., "ring the handbell"), whereas others did not (e.g., "snap your fingers"). Thirty-four of these events were employed as experimental items, and the remaining 14 events were buffer items. The events were randomly assigned to two lists, each containing 17 experimental items and 7 buffer items that served as the primacy (3 events) and recency (4 events) items in the list. Six different sets of list pairs were generated in this way, each pair containing the same 34 experimental items occupying the middle portions of the lists, and the same 14 items occupying the primacy and recency portions. The same six sets of list pairs were used for each of the three event presentation formats.

Design. Two groups of 18 subjects were randomly assigned to each of the three action-event formats. Each group was split into three subgroups, and each of the six resulting subgroups received a different set of lists.

Procedure. In the SPT condition, the subject sat at a table, half of which was screened from his/her view. The objects used in the tasks were behind the screen. The experimenter presented each task by instructing the subject what to do. If the task required the manipulation of an object, the object was presented together with the task instruction. Immediately following completion of the task, the experimenter retrieved the object and replaced it behind the screen. Subjects were instructed to perform the tasks as quickly and efficiently as possible; the mean presentation rate was approximately one task per 5 sec. Following each list the subject was to write down as many of the events as he/she could remember, describing each in two or three words.

In the EPT condition, the experimenter read aloud each task instruction and then performed the task. In the TI condition, the experimenter simply read aloud the series of instructions. Again, mean rate of presentation was approximately one event per 5 sec.

The subjects were tested individually in the SPT condition, and in groups of 2-6 in the EPT and TI conditions. In all three conditions, the subjects were allowed 120 sec for the written recall of each list.

Results and Discussion

Recall probabilities were calculated for each of the 34 nonbuffer events for each of the three presentation formats. The test-retest reliabilities are shown in Column 1 of Table 1. The SPT and EPT events showed significant test-retest correlations; however, test-retest correlation for TI just failed to reach the .05 significance level.

Recall probabilities within each format were also calculated by collapsing the data across test and retest. Corre-

Table 1
Coefficients for Test-Retest Correlations Within Each Event Format, and for Intertest Correlations Between Formats (r), in Experiment 1

	Test-Retest			Between Conditions	
	r	r_p		r	r_p
SPTs	.57	.56	SPT/EPT	.72	.71
EPTs	.67	.66	EPT/TI	.54	.52
TIs	.32	.22	SPT/TI	.45	.44

Note—The r_p coefficients were calculated by partialling out the objects/no-objects variable. In all cases, $N = 34$. SPT = subject-performed task; EPT = experimenter-performed task; TI = task instruction.

Table 2
Mean Proportion of Items Recalled Within Each Condition of the 2 × 3 × 2 Factorial in Experiment 1

	Test			Retest		
	SPT	EPT	TI	SPT	EPT	TI
Objects (n=18)	.53	.48	.32	.53	.44	.26
No Objects (n=16)	.59	.58	.49	.58	.53	.34

Note—SPT = subject-performed task; EPT = experimenter-performed task; TI = task instruction.

lations between the recall probabilities of the three classes of event, based on the collapsed data, are shown in Column 3 of Table 1. All three coefficients are significant, although the trend is as we expected; the degree of relationship was greater between the recall probabilities of events presented in the SPT and EPT formats than between those presented in the SPT and TI formats. This trend is somewhat difficult to interpret, however, because of the low test-retest reliability in the case of TIs.

Further analysis of the data was carried out using analysis of variance. For this analysis, the events were divided into those that employed objects that were temporarily introduced into the test situation for a specific SPT or EPT (e.g., a toothpick or a deck of cards) and those that did not. As already discussed, the rationale for this division was that events that did not require extrasituational objects might be favored at recall because of a greater preponderance of environmental cues.

The significance of the (fixed) experimental effects was tested using a mixed-effects analysis of variance, where the events constituted the random effects variable. The analysis employed a three-way mixed design, where test-retest and event format were within-events variables and objects/no-objects was a between-events variable. Means for the 12 conditions of the design are given in Table 2. The analysis yielded significant main effects for the test-retest variable [$F(1,32) = 7.37, MSe = 29.77$] and for format [$F(2,64) = 22.71, MSe = 8.69$]. Although the no-object events yielded higher recall than the object events, the effect of this variable failed to reach the .05 level of significance [$F(1,32) = 3.14, MSe = 225.26, p < .10$]. None of the interactions approached significance.

Although the objects/no-objects variable did not reach the .05 significance level, test-retest and between-event format correlations were recalculated partialling out this variable. As may be seen from Columns 2 and 4 in Table 1, removing the influence of the objects/no-objects variable left these correlation coefficients virtually unchanged.

In sum, then, we may conclude from Experiment 1 that some events are recalled reliably better than others, that there is a tendency for no-object events to be recalled better than object events, and that the event format plays an important role in determining recall level. In addition, the interevent correlations suggest that those characteristics of the action events that determine recall probability show some degree of generality across the event formats, especially between the SPTs and EPTs. Further discus-

Table 3
Recall Probabilities for High- and Low-Recall Events, as Measured in Experiments 1 and 2

High-Recall	Experiment 1	Experiment 2	Low-Recall	Experiment 1	Experiment 2
		Subject-Performed Tasks			
Stretch your legs*	.81	.83	Stick out your tongue	.33	.75
Put cap on pen	.78	.33	Spell COLD*	.33	.42
Roll the pen*	.75	.67	Put the top on bottle	.33	.92
Point to the door*	.72	.92	Make binoculars with		
Stand up/sit down*	.69	.92	your hands	.33	.50
Blink three times*	.69	.75	Clap your hands*	.33	.42
Push the toy car	.67	.58	Open the book*	.31	.17
Salute	.64	.50	Knock the walnuts together*	.28	.33
Throw the dice*	.64	.75	Flip the coin*	.28	.33
			Look through the		
			magnifying glass*	.25	.33
Mean	.71	.69	Mean	.31	.47
		Experimenter-Performed Tasks			
Stand up/sit down*	.86	1.00	Sharpen the pencil*	.31	.38
Blink three times	.83	.58	Make binoculars with		
Stretch legs*	.78	.75	hands*	.25	.17
Throw the dice*	.75	1.00	Scratch nose	.22	.58
Push the toy car*	.72	.83	Put top on the bottle	.22	.50
Blow a kiss*	.69	.67	Fold arms*	.19	.33
Stack checkers*	.69	.67	Fasten the pin*	.19	.33
Point to the door	.67	.50	Look through the		
Put cap on the pen	.64	.25	magnifying glass*	.17	.25
			Open the book	.14	.58
			Flip the coin*	.11	.25
Mean	.74	.70	Mean	.20	.38
		Task Instructions			
Look up at ceiling*	.58	.50	Sharpen the pencil	.19	.42
Spell COLD*	.47	.42	Stick out your tongue	.19	.33
Yawn*	.47	.38	Wave goodbye*	.17	.25
Stand up/sit down*	.44	.42	Toss the ball	.17	.42
Push the toy car	.44	.17	Put glasses in case*	.17	.15
Blink three times*	.42	.50	Open the door*	.14	.08
Put cap on pen*	.42	.67	Flip the coin*	.11	.17
Stretch your legs	.42	.25	Fasten the pin*	.08	.08
Throw the dice	.42	.33	Look through the		
			magnifying glass*	.08	.08
Mean	.45	.40	Mean	.15	.32

*Those events selected for further study.

sion at this point will be limited to the latter two conclusions.

That the SPTs and EPTs yielded higher recall levels than the TIs was not unexpected. As already discussed, recoding would be necessary to yield rich, concrete representations of the TIs in memory. The between-format differences in recall level can be interpreted as signifying that any recoding performed on the TIs did not raise their representations in memory to the level of richness or concreteness of the SPT or EPT representations.

The pattern of the between-format correlations is also explainable using the recoding notion. Thus, the correlations between the EPTs and SPTs yielded higher coefficients than did correlations involving TIs, probably because the recoding of the TIs produced memory representations that differed somewhat from those of the overt action events. Not only would we expect a recoding of the TIs to produce a less rich or concrete trace than the SPTs or EPTs, but we would also expect more across-subject variability in the TI representations. For example,

the event "push the toy car" involved the same car for all subjects participating in the SPT and EPT conditions. In the TI condition, however, recoding would presumably produce a wide variety of toy car images.

In order to optimize our chances of identifying event characteristics related to recall probability, we selected the highest nine and lowest nine events in each format for further study. These events are shown in Table 3, from which it may be observed that there is a fair amount of overlap among the events listed under the three format headings. There are seven high- and six low-recall events common to the SPT and EPT formats, six high- and five low-recall events common to the EPT and TI formats, and six high- and four low-recall events common to the SPT and TI formats.

EXPERIMENT 2

It should be noted that although six different orderings of the 34 target events were used in Experiment 1, the list contexts for each set of 18 events in Table 3 were es-

sentially limited. Before proceeding to explore possible recall-critical characteristics of the high- and low-recall events, therefore, these events were set into new list contexts and their recall remeasured.

Method

Subjects. Subjects were 54 volunteers, ranging in age from 18 to 50 years, who were visitors to the Ontario Science Centre.

Materials. Forty-eight new tasks were selected from the task pool. These tasks were divided into three 16-item lists. Into each list, 2 high- and 2 low-recall events (as determined in Experiment 1) were randomly inserted among items occupying Serial Positions 4 through 12. This gave three 20-item lists, in each of which the first 3 and the last 4 items were considered buffers. The division procedure was repeated twice more, giving three sets of three lists, where each set contained 6 high-recall and 6 low-recall items, and each of the 9 high-recall items and 9 low-recall items occurred twice over the three sets of lists.

Design and Procedure. As in Experiment 1, all three event formats were included for study. A sample of 18 subjects was tested within each format. Each sample was divided into three subgroups of 6 subjects, and each subgroup received a different set of lists. The procedure was identical to that used in Experiment 1.

Results

An analysis of variance, similar to that used in Experiment 1, was carried out on the total data in Experiment 2 (all nonbuffer events common to the three formats), using a 3 (formats) × 2 (objects/no-objects) mixed design. The mean recall probabilities for each of the six conditions of this design, given in Table 4, show the same trends as in Experiment 1. The recall of the action sentences was poorer than that of the SPTs and EPTs, and the events that used extrasituational objects were more poorly recalled than those that did not. The analysis showed both main effects to be significant [$F(2,66)$ = 11.66, MSe = 244.7, for the format variable; $F(1,33)$ = 10.81, MSe = 1,746.17, for the objects/no-objects variable]. The interaction between the two main variables was significant ($F < 1.0$).

An additional point of similarity between the data from Experiment 1 and those from Experiment 2 was in the actual levels of recall obtained. If the data in Experiment 1 are collapsed across the test–retest, the means in five of the resulting six conditions show very close agreement with their counterparts in Experiment 2 (see Table 4), despite the fact that the lists in Experiment 2 were four items shorter than those in Experiment 1. Because of this high degree of correspondence in the overall recall levels of the two experiments, it was considered justifiable to

Table 4
Mean Proportion of Items Recalled Within Each Condition of the 3 × 2 Factorial in Experiment 2

	SPT		EPT		TI	
Objects ($n=22$)	.40	(.53)	.44	(.46)	.27	(.29)
No Objects ($n=13$)	.60	(.59)	.58	(.56)	.41	(.42)

Note—The values in parenthesis were obtained in Experiment 1. SPT = subject-performed task; EPT = experimenter-performed task; TI = task instruction.

Table 5
Coefficients (r) for Interformat Correlations and r_p Coefficients Obtained by Partialling Out the Objects/No-Objects Variable in Experiment 2

	r	r_p
SPT/EPT	.62	.56
SPT/TI	.40	.26
EPT/TI	.21	.01

Note—In all cases, $N=35$. SPT = subject-performed task; EPT = experimenter-performed task; TI = task instruction.

make direct comparisons between the recall of the individual events included in both experiments (see Table 3).

There was little overall regression toward the mean for the high-recall items. There is, however, some evidence of such a regression in the case of the low-recall items. Despite this trend, it is still possible to make a distinction between most of the high- and low-recall items in Experiment 2. Since a clear distinction could not be made for all events, however, the number of items selected for further study was reduced from 18 to 12 for each event format. These items are indicated by asterisks in Table 3.

As in Experiment 1, simple linear and partial correlations were computed between the recall probabilities of those events that were common to the three formats, excluding the buffers. These are given in Table 5. The trend is again similar to that obtained in Experiment 1, with the SPT and EPT recall probabilities correlating more highly with each other than with those of the TIs. Although the objects/no-objects variable proved to be a significant determinant of recall probability in Experiment 2, the partialling out of this variable did not have much effect on the correlation coefficients, especially in the case of the SPT/EPT correlation. Consequently, although Experiments 1 and 2 suggest that one factor that subscribes to recall probability may be the presence or absence of extrasituational task objects, there is at least one other important factor involved.

We subjected the asterisked items in Table 3 to further study in Experiment 3 in an attempt to identify event characteristics correlating with recall probability.

EXPERIMENT 3

The variables included for study in Experiment 3 were selected from the word memory literature. One variable that affects word recall and that could conceivably affect the recall of action events is familiarity (Hall, 1954; Raymond, 1969; Sumby, 1963). A possible connection between this variable and action-event recall was therefore sought by having subjects rate those events marked with asterisks in Table 3 on a familiarity scale and observing whether the high-recall events were also rated high in familiarity.

A second variable that was investigated as a possible correlate of recall probability in the action events was vividness. By having subjects rate the events on this variable, we could measure whether the high-recall events also

tended to be the most vivid. (There is some evidence for considering vividness as a factor in word memory, although this is not overly convincing [Tulving, McNulty, & Ozier, 1965].)

For practical reasons, only the EPTs and TIs were studied in Experiment 3. Ratings of familiarity and vividness were obtained for both types of events. In addition, the TIs, being verbal in nature, were rated for ease of imaging (Paivio, Yuille, & Madigan, 1968).

Method

Subjects. Subjects were 96 volunteers, ranging in age from 18 to 50 years, who were visitors to the Ontario Science Centre. They were divided into four equal-sized samples, two of which rated the EPTs and two of which rated the TIs.

Design and Procedure. Each set of items was rated by two subject samples, the procedure differing slightly in the two cases.

Procedure A was as follows: In the TI condition, the experimenter read aloud the list of 12 TIs three times, using the same fixed random order each time. Subjects used a 5-point scale to rate a different variable—vividness, familiarity, and ease of imaging—on each presentation of the list. For the EPTs, only two list presentations were required, since ease of imaging was not rated.

Procedure B was similar to Procedure A, except that each list was presented only once, and the subject was required to rate each event on all the appropriate scales following its single presentation.

In both procedures, the subjects were tested in six groups of 4, each group receiving a different presentation order of the events. The ordering of the rated variables was counterbalanced across these subgroups.

Results

The mean ratings obtained from the two procedures are given in Table 6, which also shows the reliabilities of these observations as measured by a test–retest (Procedure A/Procedure B) correlation coefficient. These reliabilities were rather high in four out of the five cases and moderate in the fifth, despite the procedural differences.

As is obvious from Table 6, the difference between the high- and low-recall events is not explainable on the basis of differences in rated vividness or familiarity (or ease of imaging). Furthermore, given that the simple linear correlations between recall probability and the rated vari-

Table 6
Mean Ratings for High- and Low-Recall Items, as Measured by Procedures A and B in Experiment 3, and Test–Retest Reliabilities (r)

	Task Instructions					Experimenter-Performed Tasks				
	High-Recall		Low-Recall			High-Recall		Low-Recall		
	Mean	SD	Mean	SD	r	Mean	SD	Mean	SD	r
					Vividness					
A	3.5	.36	3.6	.48	.84	3.2	.43	3.0	.35	.57
B	3.3	.32	3.3	.49		3.2	.42	3.1	.26	
					Familiarity					
A	3.7	.46	3.6	.67	.83	3.4	.59	3.2	.83	.85
B	3.6	.65	3.5	.67		3.5	.70	3.2	.91	
					Imaging					
A	4.1	.27	4.1	.40	.75					
B	4.2	.26	4.1	.35						

Note—Maximum rating value = 5.0. SD = standard deviation.

ables proved to be very low (between 0.00 and 0.20), the possibility of relating recall level to a weighted combination of the rated variables was not pursued. Thus, although the recall probabilities of action events may be correlated with vividness or familiarity (or ease of imaging) in a general sense, this approach has clearly failed to provide an explanation for the differences in recall probability observed in the event samples under investigation.

EXPERIMENT 4

As already discussed, our action events differ from word events because the former leave traces in the physical environment that have potential value as retrieval cues. This notion has received some support in Experiments 1 and 2, in the finding that those events that involved the use of extrasituational objects tended to be less well recalled than those events that did not use such objects. Experiment 4 was designed to further test the notion that at least part of the variance in recall probability in the action-event samples depends on the differential availability of environmental cues. We tested this notion by having two recall conditions, one in which the subjects performed the recall phase of the experiment with their eyes open, and a second condition in which they recalled with their eyes shut. Thus, in one condition the physical environment was present at recall, whereas in a second condition it was virtually removed. We predicted that if the recall differences between the object and no-object events depend on differences in the cuing value of the environment, this difference should decrease when the environment was removed (eyes-shut condition).

The possibility that the environment plays a more general role in determining recall probability was also investigated in Experiment 4. Over and above the objects/no-objects comparison, we considered the possibility that the environment might provide more powerful retrieval cues for some events than for others, and thus contribute to the interevent differences in recall. More specifically, this hypothesis would predict that the eyes-shut (environment-absent) condition would have a more deleterious effect on the recall of high-recall than of low-recall events.

Method

Subjects. Subjects were 72 volunteers, ranging in age from 18 to 50 years, who were visitors to the Ontario Science Centre.

Materials. All three event formats were tested. The 12 asterisked events from each format (see Table 3) were divided into two sets, each of which contained 3 high- and 3 low-recall events. As in Experiment 2, each of these sets was randomly mixed into a list of 12 other events of appropriate format, with the constraint that the critical events should occupy Serial Positions 4 through 14 in the resulting 18-item lists. This procedure was repeated twice more, thus yielding three sets of list pairs. The buffer items used in the first set of lists were retained as buffers in the two other list sets.

Design. The subjects were divided equally among the three formats. Within each format, 12 subjects were tested under each of the two experimental treatments (eyes open and eyes closed). A further subdivision within each group gave three subgroups, each

Table 7
Mean Proportion of High- and Low-Recall Items, Recalled Under
the Open- and Closed-Eyes Conditions in Experiment 4

	High-Recall Items			Low-Recall Items		
	Eyes Open	Eyes Shut	Difference	Eyes Open	Eyes Shut	Difference
SPTs	.74	.65	.09	.46	.40	.06
EPTs	.81	.78	.03	.36	.35	.01
TIs	.52	.42	.10	.26	.18	.08

Note—SPT = subject-performed task; EPT = experimenter-performed task; TI = task instruction.

consisting of 4 subjects. Each subgroup received a different set of list pairs. The SPT subjects were tested individually; the EPT and TI subjects were tested in their subgroups.

Procedure. The subjects in the eyes-open conditions were tested using the procedure already described in Experiments 1 and 2, except that recall was verbal instead of written. The subjects in the eyes-closed conditions were instructed to shut their eyes as soon as list presentation was complete and to keep them shut until the end of the recall period. The experimenter observed the subjects closely to ensure that they followed this instruction. Recall with closed eyes was again verbal. In all conditions the subjects were allowed 90 sec for the recall of each list.

Results

Table 7 shows mean recall levels for the high- and low-recall items. Because of the relatively few items in the high- and low-recall categories, the subjects were substituted for the events as the random effects variable in the analysis of variance.

The eyes-open conditions yielded significantly better recall than did the eyes-closed conditions [$F(1,66) = 3.75$, $MSe = 1.35$]. This trend was not any greater for the high-than for the low-recall events, however; both the two-way and three-way interactions involving the high/low recall and the eyes-open/eyes-closed variables yielded $Fs < 1.0$. Consequently, although these findings suggest that the presence of the acquisition environment aids the recall of action events, they do not support an environmental cuing explanation for the recall differences between the extreme groups of events selected from Experiments 1 and 2.

An analysis of the data for all the nonbuffer items was also performed. In addition to the eyes-open/eyes-closed variable, this analysis included the objects/no-objects variable. These results are given in Table 8. A 3 (formats) × 2 (objects/no-objects) × 2 (eyes-open/eyes-closed) analysis of variance showed the objects/no-objects variable to be significant [$F(1,66) = 6.29$] and to interact significantly with format [$F(2,66) = 3.54$], $MSe = 241.74$

in both cases. The effect of the eyes-open/eyes-closed variable failed to reach the .05 level of significance in this analysis [$F(1,66) = 2.83$, $MSe = 381.52$, $p < .10$]. Furthermore, although the data in Table 8 suggest an interaction between eyes-open/eyes-closed and format, the analysis failed to support this observation [$F(2,66) = 1.74$, $MSe = 381.52$]. More importantly, although the eyes-open/eyes-closed variable appears to have had a greater effect with the no-object than with the object events, neither the interaction between these two variables [$F(1,66) = 1.92$] nor the triple interaction [$F(2,66) < 1.0$] proved to be significant. Thus the advantage of the no-object over the object events appears to be at best only weakly related to a greater preponderance of environmental cues.

EXPERIMENT 5

Since the word memory and the environmental cuing approaches produced little insight into the factors underlying the interitem differences in action-event recall, we turned to a third approach, namely an analysis of the events into action and object components. For example, the event "throw the dice" is a high-recall item and "flip the coin " is a low-recall item (see Table 3). The question we asked in Experiment 5 was whether the memorial advantage of "throwing the dice" over "flipping the coin" depends on the dice object being more memorable than the coin object, or the throwing (on the table) action being more memorable than the flipping (into the air) action, or on the unique combination between the actions and objects represented in these two action events. An answer to this question was sought in Experiment 5 through the use of recombinant events, that is, events composed of the action components taken from high-recall events and object components taken from low-recall events, and vice versa. To revert to our example, this

Table 8
Mean Proportions of Nonbuffer Items Recalled Within Each Condition of the
Eyes-Open/Eyes-Closed × Objects/No-Objects Factorial Design in Experiment 4

	Objects			No Objects		
	Eyes Open	Eyes Shut	Difference	Eyes Open	Eyes Shut	Difference
SPTs	.62	.51	.11	.70	.53	.17
EPTs	.56	.60	− .04	.58	.55	.03
TIs	.26	.28	− .01	.46	.40	.06

Note—These data are based on 22 items in each format, divided into 13 object and 9 no-object events for the SPTs and EPTs, and into 12 object and 10 no-object items for the TIs. SPT = subject-performed task; EPT = experimenter-performed task; TI = task instruction.

meant measuring recall probabilities for "throw the dice," "flip the coin," "throw the coin," and "flip the dice."

Method

Subjects. Subjects were 240 volunteers ranging in age from 18 to 50 years, who were visitors to the Ontario Science Centre.

Materials. Because of the recombinant manipulation, only a few of our action events were suitable for use in Experiment 5. These were selected from among all the nonbuffer SPTs and EPTs used in Experiments 1 and 2, to meet two criteria. First, any event selected for inclusion in Experiment 5 had to be matched with a second event on the basis that the action and object components were interchangeable. Second, of the events in each pair, one had to have a substantially higher recall value than the other in the two event formats. This selection procedure resulted in eight pairs of events, all of which satisfied the above criteria.

To aid in our description, the actions and objects belonging to the high-recall events were respectively designated HiA and HiO components, and those belonging to the low-recall events were designated LoA and LoO components. The high- and low-recall events constituting each pair could then be designated HiA-HiO and LoA-LoO events, respectively. The recombinant events, formed by switching action and object components within each matched pair, were designated HiA-LoO and LoA-HiO events.

To measure recall probabilities for the 32 events thus created, we constructed four lists. Each list consisted of 7 buffer items (3 primacy and 4 recency), 8 filler items, and 8 critical items inserted randomly into the nonbuffer portion of the list. The critical items in any one list consisted of two each of the four combination types HiA-HiO, LoA-LoO, HiA-LoO, and LoA-HiO. Each list contained a different set of the critical events, which means that each of the 32 critical component pairings occurred in only one of the four lists. One final point to note about the list construction is that each event component (action or object) occurred only once within any one list. The same lists were used for all three formats of action events.

Although the selection of events was based only on prior SPT and EPT recall probabilities, all three event formats were included in Experiment 5. It should be noted, however, that although the TI format used the same eight pairs as did the other two formats, all of the TI events did not meet the second criterion. (If this criterion had been strictly applied across all three formats, only three pairs of items would have been available for the study.) In an attempt to compensate for this, we modified the TIs to provide more precise descriptions of the action events. For example, the instruction "throw the dice" was made more specific: "throw the dice across the table."

Design. Eighty subjects were assigned to each format. Each format group was subdivided into four groups of 20 subjects, each of which received one of the four lists.

Results

The results of Experiment 5 are given in Table 9. The first three columns of data in the table give the recall probabilities values for the SPTs, EPTs, and TIs, respectively. For the purposes of analysis, the eight sets of events were regarded as eight levels of a random effects variable, and the resulting 3 (formats) × 4 (combination types) matrix was treated as a within-sets analysis of variance. As expected, the format variable proved to be highly significant [$F(2,14) = 13.80$, $MSe = 15.33$], clearly reflecting the relatively low-recall levels associated with the TI condition. The combination type also proved to be significant [$F(3,21) = 17.71$, $MSe = 7.66$], as did the interac-

Table 9
Mean Proportion of Events Recalled for the Four Action-Object Combinations Employed in Experiment 5

Action-Object Combination	Event Format			
	SPT	EPT	TI	TI (Reclassified)*
HiA-HiO	.63	.58	.21	.30
HiA-LoO	.42	.52	.21	.21
LoA-HiO	.34	.30	.19	.15
LoA-LoO	.24	.23	.13	.08
LSD	.16	.14	.13	.10

Note—SPT = subject-performed task; EPT = experimenter-performed task; TI = task instruction; LSD = least significant difference. *For an explanation of the TI (reclassified) column, see text.

tion between the two main variables [$F(6,42) = 2.99$, $MSe = 7.73$].

Inspection of the data readily identified the source of this interaction. For the SPTs and EPTs, the HiA-HiO events were recalled better than were the LoA-LoO events in all eight sets; for the TIs this was not the case. Given that the TIs were not originally chosen to conform to the same criteria as the other two formats (see Materials), this latter result was not entirely unexpected. For this reason, the TI events were reclassified such that the events with the highest and lowest recall scores within each set were designated HiA-HiO and LoA-LoO events, respectively, and the recombinant events were relabeled accordingly. The results of this procedure, given in Column 4 of Table 9, show the same pattern as the other two formats. The reclassified TI recall probabilities are more relevant to our discussion, since they conform to the same two criteria as those of the other two formats, and further discussion will deal only with these.

Least significant difference (LSD) tests applied to the combination type variable yielded similar results for each format. Thus, of the two types of recombinant events, only the HiA-LoO events yielded significantly higher recall probabilities than the LoA-LoO events, this being true for all three formats. A second finding common to the three formats was that the LoA-HiO values were significantly lower than the HiA-HiO values. Only in the case of the SPTs, however, did the mean recall probabilities for the HiA-LoO events prove to be significantly lower than those of the HiA-HiO events.

Two general conclusions may be drawn about the events used in Experiment 5. One conclusion is that the action events are analyzable into memorially meaningful components; the recall probabilities of the action events can be predicted from the recall probabilities of their components. The second general conclusion is that the action components appear to have the major role in determining recall, with the object components playing only a supporting role.

A third conclusion, which at first sight may seem somewhat surprising, is that the importance of the action component appears to be at least as great for the EPTs and

the TIs as for the SPTs. Thus observed or (presumably) imagined actions carry as much weight in determining recall probabilities as do actions actually carried out by the subjects.

GENERAL DISCUSSION

Our findings may be summarized as follows:

1. There are reliable interitem differences in the recall of action events.

2. The SPT and EPT formats appear to provide very similar findings, both as regards mean recall level and as regards the ranking of the individual events on recall probability. The task instructions yielded generally lower recall levels and showed less agreement in recall probability rankings. To use the Bäckman and Nilsson (1985) metaphor, this result suggests that TIs are not recoded to the extent that they yield equally rich memory traces as those yielded by the SPTs and EPTs. It is interesting to note, in this connection, that Bäckman and Nilsson (1984, 1985) reported no significant differences between recall levels for SPTs and for TIs in samples of young adults. This failure to demonstrate differences between SPT and TI recall levels may depend on the fact that the lists used by Bäckman and Nilsson were very short, each list containing only 12 items.

3. The approach based on our knowledge of word memory findings did not provide any insights into interevent differences in recall, the recall probabilities showing virtually no correlation with rated familiarity, vividness, and, in the case of the TIs, imageability.

4. There was a tendency for events involving objects that left no physical traces in the environment to be recalled more poorly than events that did not involve such objects. We hypothesized that the superiority of the no-object events might depend on the presence of environmental retrieval aids. Since the eyes-closed manipulation did not significantly modify the objects/no-objects differences, however, a mentally reconstructed environment must be capable of supplying as much retrieval aid as does the physical environment, if our original hypothesis is to have any validity. This supposition is not without empirical support, since the mental reinstatement of environment has been found to function as well as physical reinstatement in studies of context effects in word recall (Smith, 1979, 1984). It should be noted, however, that in Smith's studies the mental reinstatement of the environment was not done spontaneously by the subjects, but in response to instructions.

5. The component analysis approach used in Experiment 5 proved to be the most promising. In addition to providing evidence that the recall probability of an action event depends, in an additive fashion, on the recall probabilities of its components, Experiment 5 demonstrated that the action component had the major role in this regard, with the object component playing only a subsidiary role. It should be remembered, however, that we were dealing with a limited sample of events in Experi-

ment 5, and it is not difficult to conceive of other action events in which the object components might well play the major role in determining the probability of recall. Consider, for example, the probable importance of the object components in determining recall probabilities of the events "pat the toy dog" and "pat the real elephant."

The question that introduced the present series of experiments (viz., Why is it easy to remember some things but not others?) can now be answered on one level for the type of action events studied here. Events have a high probability of recall when they include an action component and, to a lesser extent an object component, which has a high probability of recall. Again this is only strictly true for the sample of events included in Experiment 5, but there is no reason to suspect that the components, and especially the action components, would not prove to be equally important in determining the recall probabilities of the other events in our pool. Unfortunately, inspection of the action components of high- and low-recall events did not reveal any clues as to the relevant variable or variables underlying the intercomponent differences; recall probability appeared to bear no relationship to such obvious variables as the amount, direction, or complexity of movement involved in the events. Future research, then, could address a reformulation of our original question, namely, Why are some actions (or objects) easier to recall than others?

A second finding that cannot be explained from the present data concerns the differences between recall of the TIs and recall of the EPTs and SPTs used in Experiment 5. Given that the recall of the TIs showed the same additive property as recall of SPTs and EPTs, with event recall probability being determined by the recall probabilities of the components, it is somewhat puzzling that the components that were Hi or Lo when presented in the verbal format were not always those that were Hi or Lo when enactment was involved.

One question that may be pursued in the present discussion, however, concerns the relationship between our action components on the one hand, and the motoric components of Saltz and Donnenwerth-Nolan (1981) and Zimmer and Engelkamp (1985) on the other. Saltz and Donnenwerth-Nolan appear to be emphasizing actual movement on the part of the subject as a prerequisite for a motoric image. Zimmer and Engelkamp, however, conceive of motoric components as involving a performance program for actions, which do not necessarily have to be overtly performed to be activated. Our action components are considered to be more in line with the Zimmer and Engelkamp concept, since these components proved to be important for determining recall probability not only for the SPTs, which involved overt movement on the part of the subjects, but also for the EPTs and TIs, which did not.

REFERENCES

BÄCKMAN, L. (1985). Further evidence for the lack of age differences on free recall of subject-performed tasks: The importance of motor action. *Human Learning*, 4, 79-87.

BÄCKMAN, L., & NILSSON, L.-G. (1984). Aging effects in free recall: An exception to the rule. *Human Learning*, 3, 53-69.

BÄCKMAN, L., & NILSSON, L.-G. (1985). Prerequisites for lack of age differences in memory performance. *Experimental Aging Research*, 11, 67-83.

COHEN, R. L. (1981). On the generality of some memory laws. *Scandinavian Journal of Psychology*, 22, 267-281.

COHEN, R. L. (1983). The effect of encoding variables on the free recall of words and action events. *Memory & Cognition*, 11, 575-582.

COHEN, R. L. (1985). On the generality of laws of memory. In L.-G Nilsson & T. Archer (Eds.), *Perspectives on learning and memory* (pp. 247-277). Hillsdale, NJ: Erlbaum.

COHEN, R. L., & BEAN, G. (1983). Memory in educable mentally retarded adults: Deficit in subject or experimenter? *Intelligence*, 7, 287-298.

COHEN, R. L., & STEWART, M. (1982). How to avoid developmental effects in free recall. *Scandinavian Journal of Psychology*, 23, 9-16.

CRAIK, F. I. M., & TULVING, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104, 268-294.

ENGELKAMP, J., & ZIMMER, H. D. (1984). Motor program information as a separable memory unit. *Psychological Research*, 46, 283-299.

HALL, J. F. (1954). Learning as a function of word frequency. *American Journal of Psychology*, 67, 138-140.

PAIVIO, A., YUILLE, J. C., & MADIGAN, S. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monographs*, 76(1, Pt. 2).

PAIVIO, A., YUILLE, J. C., & ROGERS, T. B. (1969). Noun imagery and meaningfulness in free and serial recall. *Journal of Experimental Psychology*, 79, 509-514.

RAYMOND, B. (1969). Short-term storage and long-term storage in free recall. *Journal of Verbal Learning & Verbal Behavior*, 8, 567-574.

SALTZ, E., & DONNENWERTH-NOLAN, S. (1981). Does motoric imagery facilitate memory for sentences? A selective interference test. *Journal of Verbal Learning & Verbal Behavior*, 20, 322-332.

SMITH, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning & Memory*, 5, 460-471.

SMITH, S. M. (1984). A comparison of two techniques for reducing context-dependent forgetting. *Memory & Cognition*, 12, 477-482.

SUMBY, W. H. (1963). Word frequency and the serial position effect. *Journal of Verbal Learning & Verbal Behavior*, 1, 443-450.

TULVING, E., McNULTY, J. A., & OZIER, M. (1965). Vividness of words and learning to learn in free-recall learning. *Canadian Journal of Psychology*, 19, 242-252.

ZIMMER, H. D., & ENGELKAMP, J. (1985). An attempt to distinguish between kinematic and motor components. *Acta Psychologica*, 58, 81-106.