Journal of Experimental Psychology:
Human Learning and Memory
1978, Vol. 4, No. 6, 656-665

Remembered Duration: Evidence for a Contextual-Change Hypothesis

Richard A. Block and Marjorie A. Reed
Montana State University

Two experiments used levels-of-processing tasks to investigate hypotheses on remembered duration of relatively long intervals. In Experiment 1, level of processing (shallow or deep) of presented information did not affect remembered duration, even though it had a substantial effect on memory for individual stimulus events. In Experiment 2, an interval containing different kinds of tasks (both shallow and deep processing) was remembered as being longer than one containing a single kind of task (either shallow or deep processing). Current formulations of event-memory, attentional, and informational hypotheses on remembered duration cannot easily explain these findings. However, the findings are consistent with a contextual-change hypothesis, which emphasizes memory for the overall amount of change in cognitive context during an interval. Implications regarding contextual factors in memory are discussed.

A retrospective judgment of duration of an interval must depend on memory encoding, storage, and retrieval processes. The memory processes that are involved can be affected by several different variations of information-processing tasks performed during the interval. For example, the remembered duration of an interval lengthens when a greater number of stimulus events are presented (e.g., Block, 1974, Experiment 1), when a more complex sequence of events is presented (e.g., Block, 1978, Experiment 2), when a task demanding greater selectivity of attention is performed (e.g., G. Underwood & Swain, 1973), and when a more boring task is performed (e.g., Hawkins & Tedford, 1976). Hypotheses on remembered duration have frequently been proposed in an attempt to explain an effect found when a single variable is manipulated. Some current hypotheses do not seem to be able to explain effects of a number of different variables on remembered duration. The present experiments were designed to test four kinds of hypotheses that seem to have some degree of generality and empirical support. We will refer to these as *event-memory, attentional, informational,* and *contextual-change* hypotheses.

Event-memory hypotheses propose that remembered duration is mediated by a process of covert retrieval of representations of stimulus events that occurred during the interval. A common assertion is that remembered duration is lengthened when a greater number of events are accessible (in storage and retrievable) at the time of the duration judgment. Thus, it has been proposed that remembered duration is lengthened with increases in "the multitudinousness of the memories which the time affords" (James, 1890, p. 624), "the size of the storage space" of the information *"remaining in storage"* (Ornstein, 1969, p. 104), and "the number of events stored and retained" (Block, 1974, p. 158). Attentional hypotheses (see G. Underwood, 1975; G. Underwood & Swain, 1973) place greater emphasis on attention than on memory. The most recent and most specific state-

Requests for reprints should be sent to Richard A. Block, Department of Psychology, Montana State University, Bozeman, Montana 59717.

ment of an attentional hypothesis (G. Underwood, 1975) proposes that two factors can lead to a lengthened remembered duration of an interval: increased selectivity of attention (when a task has high attentional demands) and attention to the passage of time (when a task has low attentional demands and is uninteresting). Relationships between these attentional factors and memory processes are not clear at the present time, although a process of covert retrieval may be assumed here also. Informational hypotheses, like attentional ones, emphasize characteristics of the information-processing task performed during an interval. For example, Vroon (1970) suggested that data on remembered duration could be explained by considering the amount of information presented or transmitted (measured in bits) during the interval. Vroon's data support this proposal, since he found that remembered duration is lengthened with a greater amount of information presented when overt responding is not required during an interval, but it is shortened with a greater amount of information transmitted when overt responding is required. It is not entirely clear how these informational variables affect memory processes (see Hicks, Miller, & Kinsbourne, 1976). A final kind of hypothesis to be considered here emphasizes psychological change during an interval. It has been proposed that "awareness of *change*" (James, 1890, p. 620) or "the number of changes observed" (Fraisse, 1963, p. 219) is the important factor influencing duration experiences. More recently, Block (1978) proposed a contextual-change hypothesis, which asserts that remembered duration is mediated by the remembered amount of change in cognitive context during an interval.

Although there are obvious differences among these hypotheses, it seems they all must attempt to explain the nature of the memory retrieval process or processes on which a judgment of duration must depend. Thus, one way to test these hypotheses is to investigate the effect of a variable on memory retrieval processes as well as on judgment of duration. This approach has been used by Block (1974, 1978). Some findings suggest that remembered duration is not mediated by retrieval processes involved in free recall of events, judgment of number of events, recognition of events out of context, or assignment of recognized events to the correct interval. For example, Block (1974, Experiment 1) found that increasing the number of words presented in an interval lengthened remembered duration but had no effect on number of words recalled, and Block (1978, Experiment 2) found that presenting a more complex sequence of visual patterns lengthened remembered duration but had no effect on recognition of the patterns or on assignment of recognized patterns to the correct interval. These findings can be explained by attentional, informational, and contextual-change hypotheses, but they are not as easily explained by event-memory hypotheses. However, these studies used manipulations of stimulus variables that had an effect on judgments of duration but no comparable effect on memory retrieval tasks. Additional evidence would come from a manipulation that had no effect on remembered duration but a substantial effect on memory. Also, the evidence might be more convincing and less vulnerable to alternative explanations if it were based on a manipulation of an instructional rather than a stimulus variable.

Experiment 1

Experiment 1 was designed primarily to test event-memory hypotheses by varying the level of processing of information presented during an interval (Craik & Lockhart, 1972). Experiments show that *deep* processing, such as at a semantic level, leads to better memory for individual stimulus events than does *shallow* processing, such as at a structural level (e.g., Craik & Tulving, 1975). The primary question of interest here is whether there also will be an effect on remembered duration. Event-memory hypotheses predict that an interval during which deep processing is performed will be remembered as being longer than one during which shallow processing is performed. Experiment 1 also provides a test of attentional hypotheses, provided one makes the reasonable assumption that deep processing ordinarily demands greater selectivity of attention than shallow processing. There is some evidence consistent with this assumption (see Griffith, 1976), although the evidence comes from use of experimental manipulations somewhat different from the present ones. It is probably possible to devise a shallow-processing task that demands substantial selectivity of attention; however, the shallow-processing task used in the present experiments was selected for its seemingly low attentional demands. Thus, it seems reasonable to assume that attentional hypotheses would predict that the deep-processing task we used will be remembered as longer in duration than the shallow-processing task we used.

In Experiment 1, subjects were instructed to process words to a relatively shallow, structural level during one interval and to a deeper, semantic level during another. Then they unexpectedly were asked to make a comparative judgment of duration of the two intervals, to estimate number of words presented in each, to give ratings of interest of the two tasks, to make recognition judgments, and to indicate the interval in which recognized words had occurred. The various memory judgments were used to assess the influence of processing level on memory, while the ratings of interest were used to rule out alternative explanations for the findings in terms of boredom or similar affective factors (see Hawkins & Tedford, 1976; G. Underwood, 1975).

Method

Materials and design. The pool of words that were used consisted of 20 three- to seven-letter nouns from each of four categories: four-footed animal, part of the human body, part of a building, and weapon. The 20 words from each semantic category were selected from the 30 most frequent responses in the Battig and Montague (1969) norms, avoiding any obviously polysemous words. Eight words from each category were selected and assigned randomly to each of two intervals. Then 2 of the 8 words per category per interval were assigned randomly to be typed in each of four different styles: upper- and lowercase block letters (IBM Elite 72) and upper- and lowercase italic letters (IBM Light Italic). Thus, 32 words were assigned to each interval, 8 from each semantic category and 8 in each type style. The remaining 4 words in each category appeared only as distractor items on the subsequent recognition test, with 1 word in each category assigned randomly to each of the four different type styles. A slide was made of each word by typing it on white paper and mounting it in a slide frame. The slides were ordered randomly within each series and placed in a slide tray, to be presented by a Kodak Carousel projector. Presentation of the slides was controlled by a tape recording that contained inaudible signals. The two intervals were identical in duration, delimited by signals spaced 64 sec apart on the tape. The mean exposure duration of individual slides within the two intervals was 2.0 sec (including a blank slide-change interval of .8 sec). However, exposure duration was varied randomly between 1.5 and 2.5 sec across slides in order to avoid a predictable, monotonous pace.

A six-page instruction and test booklet was assembled for each subject. Page 1 contained instructions for the task to be performed during the first interval. Instructions said that a series of slides would be shown, with one word per slide. Subjects were asked to count the number of words that were instances of one of four semantic categories or four type styles. A sample word that was an instance (but which was not otherwise used in the experiment) was given. For example, some subjects in the deep-processing condition were asked to "count the number of words that are in the category *part of the human body*, (such as) the word 'shoulder.'" Some subjects in the shallow-processing condition were asked to "count the number of words that are typed in *regular, capital (uppercase) letters*, (such as) the word 'BASEMENT.'" Immediately below was a blank line on which to write the total number of instances counted during the interval. Page 2 contained similar instructions. Half of the subjects performed the deep-processing task first (Group DS), while half performed it second (Group SD). Page 3 contained instructions for the comparative duration judgment, as well as two lines, a 50-mm line above a 100-mm line. Instructions said to delimit a line length on the 100-mm line corresponding to the apparent duration of one (comparison) interval relative to the other (standard) interval, represented by the 50-mm line. Each line was labeled either "First Series of Slides" or "Second Series of Slides." For half of the subjects (an equal number in the two groups), the first interval was the standard, while for the other half, the second was the standard. Page 4 asked all subjects to estimate the number of slides that appeared in each of the two series. Page 5 asked subjects to "indicate how interesting each of the two counting tasks were" on a 7-point scale from 1 (very uninteresting) to 7 (very interesting).

Page 6 contained combined recognition-memory and interval-discrimination (often called list-discrimination) instructions, followed by 48 randomly ordered test words, with *no, yes,* and a blank line to the right of each. Each test word appeared in the same type style to which it had been assigned originally. Only half of the presented words were tested (one randomly selected word of the two in each combination of semantic category and type style). Thus, the 48 test words included 16 from the first interval, 16 from the second, and 16 distractor words from neither interval. Instructions said to decide whether or not each word had appeared in either series of slides, circling either yes or no. Then, if a given word was recognized, subjects were asked to judge whether the word appeared in the first or the second series, writing either 1 or 2 on the line. If unsure, they were asked to guess.

Each instruction and test booklet had a unique combination of pages 1, 2, and 3. A total of 64 booklets were needed, representing all combinations of the eight different page 1 instructions, four possible page 2 instructions (constrained by page 1), and two different page 3 instructions.

Subjects. Subjects were 64 introductory psychology students, both male and female, who volunteered for the experiment. All received some class credit for participating. The data of an additional 11 subjects were discarded because of their failure to understand or follow instructions. Most of these subjects had been instructed to count words typed in upper- or lowercase "regular" letters for the shallow-processing task. However, subjects were not explicitly told that some words had been typed in italic type and that these words were not to be counted. Thus, most subjects whose data were discarded reported about 16, rather than 8, instances during the shallow-processing interval.

Procedure. Subjects participated in groups of from four to six. At the outset, subjects were told to read the instructions on page 1 of the booklet. Then the first series of slides was presented, and at the end subjects were reminded to report the number of instances counted. The same procedure was followed for the second series, which began 60 sec after the end of the first series. Subjects were given ample time to complete pages 3, 4, 5, and 6—120 sec, 60 sec, 60 sec, and 300 sec, respectively.

Results and Discussion

Table 1 shows mean performance of subjects on each dependent variable. Unless of particular interest, only significant effects are discussed, all of which are reliable beyond the .01 level.

Number of instances reported. For all subjects, the number of instances reported during each interval was close to the actual number presented. The accuracy of each subject in counting instances and not counting noninstances cannot be determined from these data. However, it seems reasonable to assume that each subject's performance was nearly, if not entirely, perfect.

Duration judgment. Each duration judgment was measured as the ratio of the apparent duration of the deep-processing interval to that of the shallow-processing interval (the *deep/shallow* ratio). Data from subjects judging the first interval were combined with data from those judging the second, since there were no significant effects of this variable. The overall mean deep/

Table 1
Mean Performance of Each Group (SD and DS) on Each Dependent Variable in Experiment 1

Processing level	Group SD	Group DS	*M*
Number of instances reported			
Deep	7.84	8.06	7.95
Shallow	8.31	8.03	8.17
M	8.08	8.05	8.06
Duration judgment			
Deep/shallow ratio	.94	1.13	1.03
Judgments of number of slides			
Deep	22.3	25.1	23.7
Shallow	24.7	22.8	23.7
M	23.5	24.0	23.7
Ratings of interest			
Deep	3.47	3.66	3.56
Shallow	3.50	3.03	3.27
M	3.48	3.34	3.41
Corrected recognition performance			
Deep	.544	.526	.535
Shallow	.257	.270	.264
M	.401	.398	.399
Corrected interval discrimination performance			
Deep	.767	.701	.734
Shallow	.651	.511	.581
M	.709	.606	.657

Note. Group SD performed shallow processing first; Group DS performed deep processing first.

shallow ratio was 1.03, a ratio not significantly different from 1.00, $t(63) = 1.14$, $SE_M = .029$. The 95% confidence interval was between .97 and 1.09. Thus, a major finding of Experiment 1 is that the level of processing of information presented during an interval has no significant effect on the remembered duration of the interval. However, the mean deep/shallow ratio was greater for Group DS than for Group SD, $F(1, 60) = 11.7$, $MS_e = .053$. This finding suggests that the first interval may have been remembered as being longer than the second. In order to test this possibility, each duration judgment was calculated as the ratio of the apparent duration of the first interval to that of the second interval (the *first/second* ratio). The overall mean first/second ratio was 1.12, a ratio significantly greater than 1.00, $t(63) = 4.27$, $SE_M = .028$. The ratio was not significantly different between Groups DS and SD, $F(1, 60) < 1$, $MS_e = .051$. These results replicate the "positive time-order error" that has been obtained previously in studies of this kind (cf. Block, 1978). A possible explanation is mentioned later.

Judgments of number of slides. The only significant effect on judgments of number of slides was a Group × Processing Level interaction, $F(1, 62) = 13.0$, $MS_e = 13.5$. The first interval was usually judged to contain more words than the second (means of 24.9 and 22.6, respectively). This effect corresponds closely with the positive time-order error in the duration judgment. It is possible that similar processes may have been involved or that these judgments simply were biased by the preceding duration judgment.

Ratings of interest. Overall mean ratings of interest were near the midpoint of the 7-point scale. There were no significant effects on the ratings. Thus, the positive time-order error in the duration judgment cannot be attributed solely to different degrees of interest during the two intervals, at least as reported by subjects later.

Recognition memory. Recognition performance was measured by using a standard correction for guessing, which is simply the hit rate minus the false-alarm rate. The false-alarm rate is the probability that a nonpresented (distractor) word was incorrectly recognized. Each subject's false-alarm rate was subtracted separately from his or her hit (correct recognition) rate for deeply processed words and for shallowly processed words. The overall mean false-alarm rate was .153. The overall mean corrected recognition score was significantly greater than zero for both deeply and shallowly processed words, $t(63) = 23.4$ and 11.5, respectively, both $SE_M = .023$. However, there was a significant effect of processing level, $F(1, 62) = 167$, $MS_e = .014$. As expected, deeply processed words were recognized more frequently than shallowly processed words. This finding, along with the finding of no significant effect of processing level on remembered duration, strongly suggests that the memory processes involved in recognition memory are different from those involved in remembered duration.

Interval discrimination. There was a tendency to assign incorrectly recognized distractor words to the shallow-processing interval. In order to correct for possible response bias, each subject's interval-discrimination performance was measured in terms of a posteriori probabilities for words from the two intervals (see Hintzman, Block, & Summers, 1973). Each probability was obtained by dividing the number of correctly recognized words that were correctly assigned to interval i by the total number of correctly recognized words that were assigned to interval i (whether correctly or not). The analysis revealed an effect of processing level, $F(1, 62) = 34.0$, $MS_e = .033$, with correct assignment of words more likely for the deep-processing interval than for the shallow-processing interval. Deep processing apparently results in superior encoding or retention of whatever memory attributes are used in interval discrimination. Performance was also better for Group SD than for Group DS, $F(1, 62) = 10.3$, $MS_e = .022$. This effect is more easily understood as an Interval × Processing Level interaction, with greater difference between deeply and shallowly processed words from the second interval. It indicates a kind of recency effect (cf. Hintzman et al., 1973), which enhances interval discrimination when deep processing is more recent but impairs it when shallow processing is last. Since a similar effect was not observed in the recognition data, interval discrimination in this experiment apparently was based on memory attributes or processes somewhat different from those involved in recognition memory (contrast with Anderson & Bower, 1972). Furthermore, it can be concluded tentatively that the memory attributes or processes involved in interval discrimination are somewhat different from those involved in remembered duration, since the pattern of results is different.

Experiment 2

The results of Experiment 1, particularly the difference between the duration judgment and recognition memory performance, are not easily explained by event-memory and attentional hypotheses on remembered duration. Informational hypotheses, however, can explain the finding of no significant effect of type of processing on remembered duration, since the amount of information transmitted was equivalent for the two levels of processing. The results are also consistent with a contextual-change hypothesis, since there was no deliberate manipulation of contextual factors. Experiment 2 was primarily designed, therefore, to test informational and contextual-change hypotheses.

A major difficulty in testing any hypothesis concerning cognitive context is the relative lack of knowledge of the salience of different kinds of contextual elements. Many different manipulations have been used in attempts to study contextual effects on memory (see, for example, Fritzen, 1977; Godden & Baddeley, 1975; Hintzman et al., 1973; Spear, 1976; B. J. Underwood, 1977). However, some manipulations produce rather small, unreliable effects (Hintzman, 1978, p. 311). This state of affairs has led one researcher, in a discussion of context, to assert that "never in the history of choice of theoretical mechanisms has one been chosen that has so little support in direct evidence" (B. J. Underwood, 1977, p. 43). Fortunately, B. J. Underwood (1977, Experiments 13, 14, and 15) investigated effects of variables thought to affect context on judgments of event recency and list identification (interval discrimination). Of relevance here is his conclusion that "process context differences can serve to establish differentiating temporal codes for memories" (p. 114). He uses the term "process context" to refer to a particular aspect of cognition, and he suggests a way of manipulating it. It is assumed that the performance of different kinds of tasks requires different cognitive processes, with a resulting change in process context. The use of levels-of-processing tasks suggests one way to contrast predictions of informational and contextual-change hypotheses.

During each interval of Experiment 2, subjects engaged in four consecutive tasks. During one interval, all of the tasks involved processing information to the same level (unmixed processing), either shallow or deep. (The unmixed-processing level was varied in order to replicate the major findings of Experiment 1.) During another interval, deep- and shallow-processing tasks were performed in alternation (mixed processing). For the same reasons as in Experiment 1, informational hypotheses predict that there will be no difference between the remembered duration of mixed- and unmixed-processing intervals. On the other hand, a contextual-change hypothesis predicts that the mixed-processing interval will be remembered as longer than either type (shallow or deep) of unmixed-processing interval. Finally, event-memory and attentional hypotheses predict that the interval containing mixed processing will be remembered as longer than the one containing unmixed-shallow processing but shorter than the one with unmixed-deep processing.

Method

Materials and design. The categorized words and type style assignments were the same as in Experiment 1. Each 32-word slide series was di-

Table 2
Mean Performance of Each Group (UM and MU) on Each Dependent Variable in Experiment 2

Processing type	Group UM	Group MU	M
Number of instances reported			
Mixed	7.94	8.16	8.05
Unmixed	8.44	7.75	8.09
M	8.19	7.95	8.07
Duration judgment			
Mixed/unmixed ratio	1.00	1.24	1.12
Judgments of number of slides			
Mixed	27.2	30.6	28.9
Unmixed	31.0	28.6	29.8
M	29.1	29.6	29.3
Ratings of interest			
Mixed	3.34	3.63	3.48
Unmixed	2.84	3.94	3.39
M	3.09	3.78	3.44
Corrected recognition performance			
Mixed	.365	.268	.316
Unmixed	.271	.328	.300
M	.318	.298	.308
Corrected interval discrimination performance			
Mixed	.603	.568	.586
Unmixed	.467	.602	.535
M	.535	.585	.560

Note. Group UM performed unmixed processing first; Group MU performed mixed processing first.

vided into four blocks of 8 randomly ordered words, with each semantic category and type style represented twice in each block. In addition to the 32 single-word slides, each series contained four instruction slides. An instruction slide preceded each block (positions 1, 10, 19, and 28), naming one of the four semantic categories (e.g., PART OF A BUILDING) or four type styles (e.g., CAPITAL ITALIC TYPE). Type of processing was manipulated by the instruction slides. In one condition, all four instruction slides named semantic categories (unmixed-deep processing) or type styles (unmixed-shallow processing) in random order. In the other condition, the four instruction slides alternated between naming a random semantic category and a random type style (mixed processing). Half of the mixed-processing intervals began with deep processing, and half began with shallow; the data were collapsed across this variable. All subjects received one mixed-processing and one unmixed-processing condition. Subjects in Group UM performed the unmixed processing first, while those in Group MU performed the mixed processing first. Thus, the design was essentially a $2 \times 2 \times 2$ factorial, with processing type (mixed or unmixed) varied within subjects and both unmixed task (shallow or deep) and task order (UM or MU) varied between subjects. Each interval was 80 sec long. Instruction slides were presented for 4.0 sec each; single-word slides were presented for a mean of 2.0 sec each, varying randomly between 1.5 and 2.5 sec as in Experiment 1.

The six-page test booklet was the same as in Experiment 1 except that it did not contain instructions on the first two pages. The interval used as the standard for the duration judgment was varied orthogonally, as in Experiment 1. Sixteen groups were required to represent all combinations of the variables.

Subjects. A total of 64 subjects were obtained in the same manner as in Experiment 1. Data from an additional 6 subjects were discarded because of their failure to understand or follow instructions.

Procedure. The procedure was similar to that of Experiment 1, with subjects participating in small groups. At the outset, they were told that they would see a series of slides, with one word per slide. They also were told that the first slide and several others in the series would be instruction slides, naming one of four semantic categories or type styles. Sample instruction slides and an instance of each were shown. Subjects were told to keep a cumulative count of the instances of each instruction slide that appeared in the single-word slides that followed it and that preceded the next instruction slide. Similar, but shortened, instructions preceded the second series. The procedure generally was similar to that of Experiment 1, with the addition of paraphrased instructions given for each page of the test booklet.

Results and Discussion

Table 2 shows mean performance of subjects on each dependent variable. Level of processing during the unmixed interval produced only two significant effects (discussed below), so data are collapsed over this variable. Because of the large number of significance tests, we decided to use the .01 level of significance in order to minimize the overall Type I error rate. Unless of particular interest, only results reliable beyond that level are discussed.

Number of instances reported. As in Experiment 1, subjects were very accurate in the number of instances reported. Largely as a result of the small variance, the Group × Processing Type interaction was significant, $F(1, 60) = 8.12$, $MS_e = .924$. The mean number of instances reported was slightly greater for the first interval than for the second (8.30 and 7.85, respectively).

Duration judgment. Each duration judgment was measured as the ratio of the apparent duration of the mixed-processing interval to that of the unmixed-processing interval (the *mixed/unmixed* ratio). Data from subjects judging the first interval were combined with data from those judging the second. The overall mean mixed/unmixed ratio was 1.12, a ratio significantly greater than 1.00, $t(63) = 3.27$, $SE_M = .036$. The 95% confidence interval was between 1.05 and 1.19. Thus, a major finding of Experiment 2 is that changes in the type of information processing performed during an interval lengthen the remembered duration of the interval. As in Experiment 1, processing level during the unmixed interval had no significant effect on remembered duration. In fact, the mean mixed/unmixed ratio was 1.12 for both unmixed-shallow and unmixed-deep conditions. However, the mean mixed/unmixed ratio was significantly greater for Group MU than Group UM, $F(1, 56) = 13.9$, $MS_e = .066$. The most reasonable explanation for these results is that a positive time-order error of similar magnitude to that in Experiment 1 is added to the effect of mixed processing in Group MU but subtracted from the effect in Group UM. Support for this explanation comes from the finding of an overall mean first/second ratio of 1.14, a ratio significantly greater than 1.00, $t(63) = 4.11$, $SE_M = .035$, and similar to the first/second ratio of 1.12 found in Experiment 1.

Judgments of number of slides. The only significant effect on judgments of number of slides was a Group × Processing Type interaction, $F(1, 60) = 7.42$, $MS_e = 27.3$. As in Experiment 1, the first interval usually was judged to contain more slides than the second (means of 30.8 and 27.9, respectively). The effect is similar to the positive time-order error in the duration judgment. However, there was no effect of processing type on judgments of number of slides, so somewhat different memory processes were involved in the two kinds of judgments.

Ratings of interest. As in Experiment 1, overall mean ratings of interest were near the midpoint of the 7-point scale. The only significant effect was a Group × Processing Type interaction, $F(1, 60) = 7.52$, $MS_e = .702$. The first task was rated slightly less interesting than the second (means of 3.23 and 3.64 on the 7-point scale, respectively). A possible explanation for this finding is that subjects were more confident in what was expected of them during the second interval. If one adopts the usual interest or boredom explanation (see Hawkins & Tedford, 1976; G. Underwood, 1975), the positive time-order error in the duration judgment in this experiment might be related to differential interest during the two intervals. However, the positive time-order error in Experiment 1 was not accompanied by differences in ratings of interest. Furthermore, the effect of processing type on the duration judgment in Experiment 2 cannot be explained by differential interest in the tasks, since processing type did not affect ratings of interest.

Recognition memory. Recognition performance was measured as in Experiment 1. The overall mean false-alarm rate was .293. Corrected recognition scores for both mixed-processing and unmixed-processing intervals were significantly greater than zero, $t(63) = 13.4$ and 11.3, $SE_M = 0.24$ and .027, respectively. There was a significant Group × Processing Type interaction, $F(1, 60) = 14.4$, $MS_e = .056$. Words from the second interval were recognized more frequently than those from the first (means of .347 and .270, respectively), a recency effect. There was also a Processing Type × Unmixed Task interaction, $F(1,60) = 38.5$, $MS_e = .056$. This interaction replicates the recognition results of Experiment 1. Words from an unmixed-deep interval were recognized most frequently (mean of .415), while words from an unmixed-shallow interval were recognized least frequently (mean of .186). Words from the mixed-processing interval showed intermediate performance, with deeply processed words recognized more frequently than shallowly processed words (means of .468 and .167, respectively). As in Experiment 1, this pattern of results differs from that obtained on the duration judgment.

Interval discrimination. There was a tendency to assign incorrectly recognized distractor words to the first interval, so a posteriori probabilities were computed as in Experiment 1. There was a significant Group × Processing Type interaction, $F(1, 60) = 13.3$, $MS_e = .017$. Correct assignment of words was more likely for the second interval than for the first (means of .603 and .518, respectively), a recency effect similar to that found in the recognition data. There was also a similar Processing Type × Unmixed Task interaction, $F(1, 60) = 26.8$, $MS_e = .017$. Correct assignment of words was more likely for an unmixed-deep interval than for an unmixed-shallow interval (means of .611 and .459, respectively). In this experiment, it seems that recognition and interval-discrimination judgments were based on similar or identical memory attributes and processes. Both measures show a pattern of results different from the pattern observed on the duration judgment.

General Discussion

A major finding of Experiment 1, which was also replicated in Experiment 2, is that the level of processing of information presented during an interval has little or no effect on remembered duration, even though it has a substantial effect on memory for individual events from the interval. Event-memory hypotheses, which predict that an interval containing processing at a deeper level would be remembered as longer, are refuted by this evidence. Attentional hypotheses are also rejected if one makes the reasonable assumption that the deep-processing task demanded greater selectivity of attention than the shallow-processing task. Experiment 2 contrasted predictions of two other hypotheses that could explain the results of Experiment 1, informational and contextual-change hypotheses. A major finding of Experiment 2 is that an interval containing different kinds of levels-of-processing tasks is remembered as longer than an interval containing only one kind of task. The amount of information presented and, presumably, transmitted was equivalent for the two types of intervals. Thus, informational hypotheses, which predict no effect of this manipulation on remembered duration, cannot explain the finding. The effect on remembered duration, however, is predicted by a contextual-change hypothesis. It assumes that different kinds of tasks require somewhat different cognitive processes, with a resulting change in a particular aspect of cognitive context, process context.

A contextual-change hypothesis can also explain other aspects of the present results. The positive time-order error—remembering the first of two equal intervals as longer—can be explained by the notion that some elements of the cognitive context change rapidly near the start of an experiment (see Hintzman et al., 1973). On the question of what memory retrieval processes mediate remembered duration, past experiments (Block, 1974, 1978), as well as the present ones, show that the processes involved in retrieving representations of individual stimulus events are somewhat different from those involved in remembering the duration of an entire interval. It also appears that the processes involved in interval discrimination—remembering when particular events occurred—are not the same as those involved in remembered duration, although it is, of course, possible that some of the same memory attributes are involved in both. Since a contextual-change hypothesis does not emphasize memory for individual stimulus events but rather memory for the overall change in cognitive context during an interval, these findings are expected.

Of course, there are lingering difficulties with any hypothesis concerning cognitive context. A major problem arises when one considers the variety of different factors typically proposed to affect cognitive context (see Anderson & Bower, 1972; Bower, 1972; Hintzman et al., 1973). These factors are thought to include conspicuous internal and external stimuli, characteristics of the information presented or the task performed, and miscellaneous cognitive and affective reactions. It is difficult to reject contextual explanations, since a manipulation of a particular factor which did not have the expected effect could be dismissed as having a relatively minor influence on cognitive context. Adequate tests of any kind of contextual explanation may occur only when more is known about the kinds of factors that are important in particular situations (cf. B. J. Underwood, 1969). On the other hand, since most variables that affect remembered duration have been mentioned as possible contextual factors, a contextual-change hypothesis is both integrative and parsimonious. If a contextual-change hypothesis on remembered duration is not rejected by future experiments, other kinds of contextual hypotheses may receive needed support. Retrospective judgment of duration may serve, then, as an index of the overall amount of change in cognitive context during an interval.

References

Anderson, J. R., & Bower, G. H. Recognition and retrieval processes in free recall. *Psychological Review,* 1972, *79,* 97–123.

Battig, W. F., & Montague, W. E. Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. *Journal of Experimental Psychology Monograph,* 1969, *80*(3, Pt. 2).

Block, R. A. Memory and the experience of duration in retrospect. *Memory & Cognition,* 1974, *2,* 153–160.

Block, R. A. Remembered duration: Effects of event and sequence complexity. *Memory & Cognition,* 1978, *6,* 320–326.

Bower, G. H. Stimulus-sampling theory of encoding variability. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory.* New York: Winston, 1972.

Craik, F. I. M., & Lockhart, R. S. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior,* 1972, *11,* 671–684.

Craik, F. I. M., & Tulving, E. Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General,* 1975, *104,* 268–294.

Fraisse, P. *The psychology of time.* New York: Harper & Row, 1963.

Fritzen, J. D. Orienting tasks as sources of contextual information in list discrimination. *Journal of Experimental Psychology: Human Learning and Memory,* 1977, *3,* 119–128.

Godden, D. R., & Baddeley, A. D. Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology,* 1975, *66,* 325–331.

Griffith, D. The attentional demands of mnemonic control processes. *Memory & Cognition,* 1976, *4,* 103–108.

Hawkins, M. F., & Tedford, W. H., Jr. Effects of interest and relatedness on estimated duration of verbal material. *Bulletin of the Psychonomic Society,* 1976, *8,* 301–302.

Hicks, R. E., Miller, G. W., & Kinsbourne, M. Prospective and retrospective judgments of time as a function of amount of information processed. *American Journal of Psychology,* 1976, *89,* 719–730.

Hintzman, D. L. *The psychology of learning and memory.* San Francisco: Freeman, 1978.

Hintzman, D. L., Block, R. A., & Summers, J. J. Contextual associations and memory for serial position. *Journal of Experimental Psychology,* 1973, *97,* 220–229.

James, W. *The principles of psychology* (Vol. 1). New York: Henry Holt, 1890.

Ornstein, R. E. *On the experience of time.* Harmondsworth, England: Penguin, 1969.

Spear, N. E. Retrieval of memories: A psychobiological approach. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 4). Hillsdale, N.J.: Erlbaum, 1976.

Underwood, B. J. Attributes of memory. *Psychological Review,* 1969, *76,* 559–573.

Underwood, B. J. *Temporal codes for memories: Issues and problems.* Hillsdale, N.J.: Erlbaum, 1977.

Underwood, G. Attention and the perception of duration during encoding and retrieval. *Perception,* 1975, *4,* 291–296.

Underwood, G., & Swain, R. A. Selectivity of attention and the perception of duration. *Perception,* 1973, *2,* 101–105.

Vroon, P. A. Effects of presented and processed information on duration experience. *Acta Psychologica,* 1970, *34,* 115–121.

Received February 3, 1978
Revision received June 1, 1978 ■