Journal of Experimental Psychology:
Learning, Memory, and Cognition
1982, Vol. 8, No. 6, 530–544

Copyright 1982 by the American Psychological Association, Inc.
0278-7393/82/0806-0530$00.75

Temporal Judgments and Contextual Change

Richard A. Block
Montana State University

Three experiments investigated effects of environmental context on temporal memory judgments. An equal number of items occurred within each of two equal durations (D_1 and D_2). Subjects subsequently were asked to judge the length of a given duration in comparison with the other, then to discriminate the correct list and serial position for each recognized item on a test. If environmental context was unchanged, D_1 was remembered as being longer than D_2; if the context was disrupted during the interval separating D_1 and D_2, this effect was reduced; and if the context prevailing during D_2 also was changed, the effect was eliminated. List discrimination was improved only if the context was changed. Serial-position judgment was not affected by either manipulation. Changes in process context—the internal context produced by the performance of specific cognitive processes—lengthened remembered duration, but the effect did not simply add on to the effects of environmental context. Results are discussed in terms of a contextual-change hypothesis.

The concept of changes in cognitive context, as it relates to processes of learning, memory, and cognition, has a long and undignified history. In an early review of experiments on changes in context, McGeoch (1942) concluded that "altered stimulating conditions" (i.e., contextual change) was a "fundamental condition in the [then-]current psychological account of forgetting" (p. 501). More recently, Underwood (1977) remarked that "no single concept is so widely used in theories of memory functioning as is the concept of context" (p. 43). The concept has played an important role in interference, stimulus sampling, multicomponent, encoding variability, and other associative theories of memory (see Crowder, 1976; Zechmeister & Nyberg, 1982). However, Underwood (1977) claimed that "never in the history of choice of theoretical mechanisms has one been chosen that has so little support in direct evidence" (p. 43).

Is the difficulty with the concept attributable to the lack of reliable effects, as Underwood implied? A perusal of the literature reveals many replicable contextual-change effects (Hewitt, 1979; Smith, Glenberg, & Bjork, 1978). Is the difficulty with the concept attributable to the vagueness of the concept, which includes many subcategories and can lead to logical and empirical paradoxes (Block & Reed, 1978; Hintzman, Block, & Summers, 1973; Underwood, 1977)? Perhaps in part the answer is yes, although ways around these pitfalls have been suggested. It seems to me that much of the difficulty with the concept of cognitive context is attributable to the almost total neglect of two issues. One issue involves determining what kinds of memory retrieval and judgment are affected by the various kinds of contextual change. The second issue involves measuring transient changes in cognitive context.

Two recent exceptions are notable. First, the work of Smith (1979; Smith et al., 1978) concerning effects of environmental-context changes on memory retrieval is an exemplary exploration of the first issue. His reliable finding that a reinstatement of environmental context improved recall, but not recognition,

This research was supported by National Science Foundation Grant ISP-8011449.

I wish to acknowledge the able assistance of several students in conducting the various experiments: Kelley Barnard (Exp. 3), Michael Campbell (Exp. 1 and 3), Rebecca Dahl (Exp. 1), Marc Griffey (Exp. 2 and 3), and Janet Martin and John Saggau (Exp. 2). Michael Campbell and John Saggau also helped to analyze the data. Special thanks are due Paul Willis for advice and assistance on statistical and computing matters. I also wish to thank Steven Smith, Kirk Smith, and two anonymous reviewers for helpful comments on an earlier draft of the manuscript.

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

may go a long way in resolving perceived problems of replicability. We cannot merely claim that memory is, or is not, improved under certain contextual conditions. We must attempt to reveal which kinds of memory retrieval and judgment are enhanced, impaired, or unaffected by a given kind of contextual change. Second, the work in my laboratory concerning contextual effects on temporal memory judgments is an exploration of both issues. The general methodology in these studies is to present two equal durations, each containing several events, and then unexpectedly ask subjects to make various temporal memory judgments. The strategy attempts to determine whether or not different kinds of temporal judgment are mediated by similar aspects of memory. The term, *aspects of memory*, is used here to refer both to encoded contextual elements and to retrieval and decision processes used in making such judgments. The strategy also attempts to reveal in what way any particular kind of judgment ultimately might provide a useful index of transient changes in cognitive context.

Some evidence supports a contextual-change hypothesis on duration judgment, which says that "retrospective judgment of duration may serve . . . as an index of the overall amount of change in cognitive context" during a time period (Block & Reed, 1978, p. 665; see also Block, 1979; Block, George, & Reed, 1980). Block and Reed studied effects of an aspect of context called process context. *Process context* refers to the internal context produced by the performance of specific kinds of cognitive processes (see Underwood, 1977, p. 100). They found that changes in process context during a time period lengthened the remembered duration of it. Another finding in Block and Reed's and other recent experiments (e.g., Block, 1978) is that of a positive time-order error: All other factors equal, or counterbalanced, the first of two equal durations (D_1) is judged to be longer than the second (D_2). The contextual-change hypothesis explains the finding by assuming that cognitive context changes rapidly near the beginning of a new experience (cf. Hintzman & Block, 1971). Thus, cognitive context might ordinarily change more during D_1 than during D_2. Additional evidence is needed to support the ultimate potential use of duration judgment as an index of contexual change.

The present experiments test the generality of the contextual-change hypothesis by using different kinds of contextual manipulations and by varying contextual change between two durations as well as within a duration. Manipulations of environmental context similar to those of Smith et al. (1978) were used. Following a comparative-duration judgment, subjects also made combined judgments of recognition, list discrimination, and serial position. The major question is whether a given kind of temporal memory judgment is affected by a change in environmental context prevailing during D_2 (see Experiments 1, 2, and 3), by a disruption in environmental context occurring during the interval separating D_1 and D_2 (see Experiment 2), by both kinds of change, or by neither. Experiment 3 concerns a somewhat different issue: Are effects of changes in environmental and process context additive?

Experiment 1

Effects of contextual changes can be assessed by comparing temporal memory judgments in two conditions. In one condition, environmental context is the same during D_2 as it was during D_1. In another, environmental context is different during D_2. Strand (1970) noted that earlier experiments on environmental context confounded changes in environmental context with psychological disruption and physical activity. Subjects in a changed-context condition must be disrupted between D_1 and D_2. In Experiment 1 the potential source of confounding is eliminated by also disrupting subjects in the unchanged-context condition. Thus, two main environmental-context conditions are compared, a disruption-only and a disruption-and-change condition. Experiment 1 also investigates possible effects of environmental context prevailing at the time the temporal judgments are made.

Method

Environmental contexts. Two rooms were selected and equipped so that both the amount and the quality of available contextual information differed as much as possible. Context C_1 was a large classroom with several doors, a linoleum floor, a large chalkboard, several tables, many desk chairs, and various objects. The room appeared cluttered. A slide projector was used to display the stimulus materials on a large screen, and the room was dimly illuminated. Subjects sat in desk chairs in a semicircle about 4 m from the screen. The experimenter was a casually dressed male. Context C_2 was a small room with one door, a carpeted floor, bare white walls, and no furniture. The room appeared uncluttered. The only noticeable equipment was a small podium, clipboards, and pencils. Large index cards were used to display the stimulus materials, and the room was brightly illuminated. Subjects sat on the floor, along the walls, about 2.5 m from the experimenter. The experimenter was a female who wore a white laboratory coat over her clothes. Neither room contained a clock; both experimenters used the same concealed digital stopwatch to time the durations. The two rooms were entered through a common hallway; the entrance doors were separated by a distance of about 15 m.

Materials. A total of 35 common occupations was selected from the Battig and Montague (1969) norms. No word contained more than 10 letters, and no occupational labels were synonymous. For C_1, each occupation was typed on plain white paper and mounted in a slide frame. For C_2, each occupation was printed in large letters on an index card. A total of 30 occupations was randomly selected and ordered, and 15 were assigned to each of two lists; the remaining 5 occupations were used only as distractor items on the memory test. A pencil and a small notepad were used by each subject for the occupation-rating task.

Two versions of a duration-judgment form were used. Each contained instructions for making a 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 in order to indicate the apparent duration of one time period relative to that of the other, represented by the 50-mm line. Each was defined as beginning when the experimenter had said, "*This* is the first [occupation]," and as ending when the experimenter had said, "*That* was the final one." Each line was labeled either "First Series of Words" or "Second Series of Words." Half of the subjects, and an equal number in each variation of experimental conditions, judged D_1 relative to D_2, while the other half judged D_2 relative to D_1.

The same memory-judgment form was used by all subjects. On the top half of the sheet were instructions, and on the bottom half were 35 randomly ordered occupations. The instructions said to decide first whether or not each occupation had been presented previously. If not, the subject was to write a 0 after the item. If he or she did remember the occupation, the instructions said to decide whether it occurred within D_1 or D_2 and then to decide its relative position in the series. Each list and position judgment was made by writing a number from 1 to 10. A scale illustrated the assignment of numbers: Each of two horizontal line segments was divided into five equal sections; the sections were labeled with the numbers 1 to 5 for occupations from the start to the end of D_1 and with the numbers 6 to 10 for those from the start to the end of D_2. Instructions said to guess if uncertain about the relative position of a particular occupation.

Subjects and design. Subjects were 112 introductory psychology students, both male and female, who volunteered for the experiment. Data of an additional 7 subjects were discarded because of their failure to follow instructions. Subjects typically participated in groups of 7 or 8, although occasionally the number varied because of the failure of some scheduled subjects to arrive for the experiment. Each group was assigned to one of eight variations in condition; a total of 14 subjects served in each variation. The variations were an orthogonal combination of two D_1 contexts, two D_2 contexts, and two judgment contexts. In other words, there were two D_1–D_2 environmental-context conditions: disruption only (C_1–C_1 or C_2–C_2) and disruption and change (C_1–C_2 or C_2–C_1). In the disruption-only condition, the judgment context was either the same as (e.g., C_1–C_1–C_1) or different from (e.g., C_1–C_1–C_2) that of D_1 and D_2. In the disruption-and-change condition, the judgment context was either the same as the D_1 context (e.g., C_1–C_2–C_1) or the same as D_2 (e.g., C_1–C_2–C_2). The variations were run in a counterbalanced order on each of 2 days. Between days, the assignment of occupation lists to D_1 and D_2 was rotated, and the order of the occupations within each list was randomized. Each group of subjects included an approximately equal number receiving either version of the duration-judgment form; across days, an equal number in each variation received each version.

Procedure. Subjects assembled in a room one floor below C_1 and C_2. The experimenter assigned to D_1 led them to the appropriate room, where each was seated and given a notepad and a pencil. The instructions and general procedure were identical for the two durations. Subjects were told that the task involved "rating how well certain occupations fit you personally," and were asked to rate each on a scale from 1 (fits you very well) to 5 (fits you very poorly). They were asked to write each rating on a separate sheet of the notepad. The instructions lasted 60 sec. The experimenter then said, "*This* is the first occupation," simultaneously starting the concealed stopwatch and presenting the first occupation. Occupations were presented at a 10-sec rate either by triggering the slide-change mechanism (in C_1) or by holding up another index card (in C_2). At the offset of the last occupation, the experimenter said, "*That* was the final occupation." Thus, each experimental duration, delimited by the articulation of the words *this* and *that*, was 150 sec. Subjects in each condition were then asked to wait in the hallway "until the next part of the experiment is set up." When 80 sec had elapsed since the end of D_1, the appropriate experimenter led subjects into the appropriate room. Including the time taken to give the same instructions, D_2 began 180 sec after the end of D_1. At the end of D_2, subjects again waited in the hallway. The appropriate experimenter called them in 80 sec later, distributed the duration-judgment form, and told them to read and follow the printed instructions. About 60 sec later the experimenter distributed the memory-judgment form. Subjects were given ample time to make the judgments.

Results and Discussion

Duration judgment. Each duration judgment was measured to the nearest millimeter (see Appendix for a subsidiary analysis of

Table 1
Mean D_1/D_2 Ratio in Each Combination of Environmental Context, Judgment Context, and Judged Duration in Experiment 1

Environmental- and judgment-context condition	Judged duration D_1	D_2	M
Disruption only			
Same context	1.04	1.17	1.10
Different context	1.20	1.09	1.14
M	1.12	1.13	1.12
Distruption and change			
D_1 context	1.20	.99	1.10
D_2 context	.94	1.01	.97
M	1.07	1.00	1.03

Note. D_1 = first duration; D_2 = second duration.

these raw data). Then each was expressed as a ratio of the remembered duration of D_1 to that of D_2 (D_1/D_2 ratio). The mean ratio in each condition is shown in Table 1. The data were first subjected to a 2 × 2 (Environmental Context × Judged Duration) analysis of variance (ANOVA) and several planned *t* tests. The analysis reveals a major finding: Mean D_1/D_2 ratio is greater if environmental context was only disrupted between D_1 and D_2 than if it was both disrupted and changed, $F(1, 108) = 4.78$, $p < .05$, $MS_e = .05$. In the disruption-only condition, judgments show a positive time-order error, because the mean ratio of 1.12 is greater than 1.00, $t(55) = 4.90$, $p < .001$, $SE = .03$. However, in the disruption-and-change condition, a time-order error is not found; the mean ratio of 1.03 is not significantly different from 1.00, $t(55) = 1.07$, $SE = .03$. Neither the main effect nor the interaction involving judged duration is significant.

Additional analyses show no significant effect of judgment context in either condition. Thus, the finding of a positive time-order error in the disruption-only condition does not depend on whether the judgment was made in the same context as D_1 and D_2 or in a different context, and the finding of no time-order error in the disruption-and-change condition does not depend on whether the judgment was made in the D_1 or the D_2 context. The interactions between judgment context and judged duration are only marginally significant.

List-discrimination judgments. Because there were only two durations, a simple list-discrimination index can be used. Considering each subject and serial-position block separately, it is the probability that an item was judged (J) to have occurred within D_1—$P(J_1)$—given recognition of the item. The recognition rate did not vary significantly across conditions; hit rates were very high and false-alarm rates were low (see Appendix). An ANOVA using planned comparisons tested for an overall linear trend across position blocks in D_1, a trend across blocks in D_2, an interaction of D_1 and D_2 trends, and a difference between D_1 and D_2 means, as well as for between-conditions interactions involving each comparison. The .01 level was used in order to minimize the Type I error rate. Unless of interest, only results reliable beyond that level are mentioned. Marginally significant findings, that is, those for which $.01 < p < .05$, are mentioned only if they are replicated in at least two experiments.

Figure 1 shows mean $P(J_1)$ given recognition in the two main conditions. Consider first the overall trends. Mean $P(J_1)$ is greater

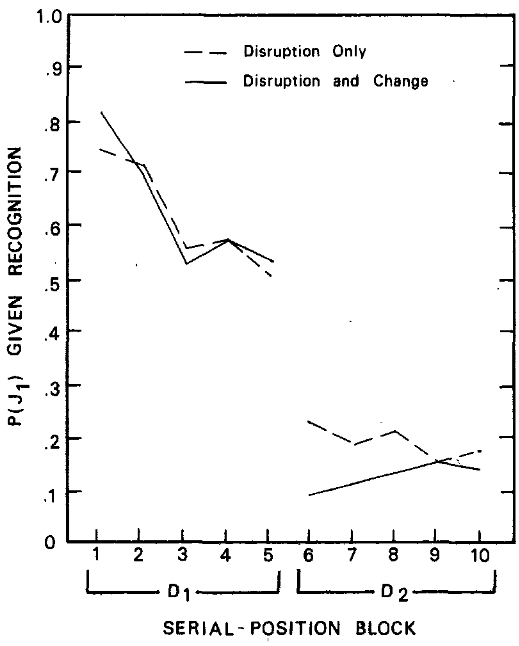

Figure 1. Mean $P(J_1)$ given recognition—probability (P) that an item was judged (J) to have occurred within D_1 given recognition—as a function of serial-position block in each environmental-context condition in Experiment 1. (D_1 = first duration; D_2 = second duration.)

for D_1 than for D_2 items, $F(1, 108) = 429$, $p < .001$, $MS_e = .14$, indicating considerable list-discrimination ability. There is a decreasing trend across D_1, $F(1, 108) = 48.9$, $p < .001$, $MS_e = .09$, but no trend across D_2 ($F < 1$); thus, there is an interaction of D_1 and D_2 trends, $F(1, 108) = 34.1$, $p < .001$, $MS_e = .06$. The only between-condition effect is a marginal difference on D_2 items, $F(1, 108) = 4.95$, $p = .03$, $MS_e = .12$. Items from D_2 are slightly less likely to be assigned incorrectly to D_1 if environmental context was changed than if it was only disrupted.

Serial-position judgments. The slopes of the position-judgment curves were measured conditional upon either a correct or an incorrect list judgment. Not all subjects contributed a datum of each kind at each serial-position block, so data from all subjects in each condition were pooled for this analysis.[1] Figure 2 shows the mean judgment of each kind in the two main conditions. One-tailed t tests assessed whether or not each slope is greater than zero, as well as whether or not each corresponding pair of slopes is significantly different.

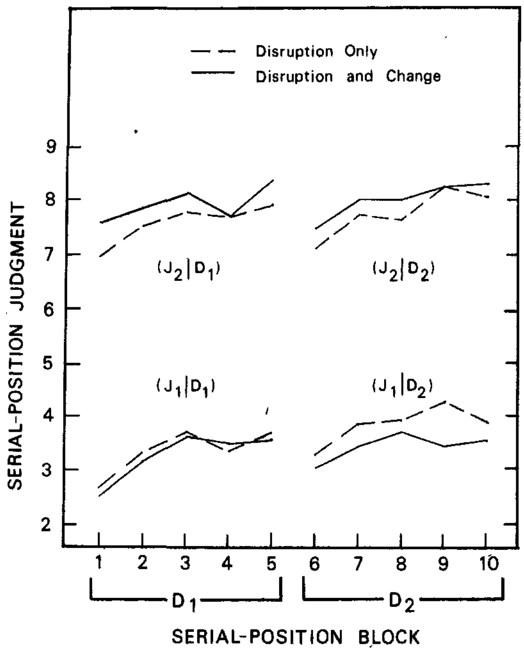

Figure 2. Mean serial-position judgment of each kind as a function of serial-position block in each environmental-context condition in Experiment 1. ($J_1|D_1$ = correct list judgment for first-duration item. $J_2|D_2$ = correct list judgment for second-duration item. $J_2|D_1$ = incorrect list judgment for first-duration item. $J_1|D_2$ = incorrect list judgment for second-duration item.)

Consider first the slopes when data from all subjects are combined. Given a correct list judgment of D_1 items ($J_1|D_1$), the slope of .24 is greater than zero, $t(3) = 2.46$, $p < .05$; and given a correct list judgment of D_2 items ($J_2|D_2$), the slope of .26 is also greater than zero, $t(3) = 5.59$, $p < .01$. Given an incorrect judgment of D_1 items ($J_2|D_1$), the slope of .19 is greater than zero, $t(3) = 2.67$, $p < .05$; but given an incorrect judgment of D_2 items ($J_1|D_2$), the slope of .11 is not, $t(3) = 1.96$. These findings show that serial-position information was available, at least if the list was correctly identified.

There are no significant differences between corresponding slopes in the two conditions, all $t(6) < .91$, and the difference between the $J_1|D_1$ and $J_2|D_2$ slopes is not significant in either condition, both $t(6) < .70$. Furthermore, given a correct list judgment, the D_1/D_2 slope ratio is .88 in the disruption-only and 1.29 in the disruption-and-change condition. This pattern of ratios is clearly different from that of the D_1/D_2 duration-judgment ratios.

Summary. The results of Experiment 1 reveal that D_1 is remembered as being longer than D_2 (a positive time-order error) if there is no change in environmental context, but D_1 is not remembered as being longer than D_2 if there is a change. Changing the environmental context also may slightly improve list discrimination on items from D_2, but there is no evidence that serial-position judgment is affected. Finally, the environmental context in which all of these kinds of temporal memory judgments are made apparently has little or no effect on the temporal judgments.

Experiment 2

Experiment 2 attempts to replicate and extend the main findings of Experiment 1. The disruption-only and disruption-and-change conditions are the same as before.

[1] The outcomes are similar in all experiments if each subject is allowed to contribute at most only a mean of each kind of observation at each serial-position block.

Because there were only a few slight effects of judgment context in Experiment 1, to simplify the design of Experiment 2 this variable is not manipulated; all subjects made judgments in a third context. However, another condition is added. If Strand's (1970) analysis is viewed in a slightly different way, the disruption manipulation obviously involves changes in environmental context during the interval separating the two durations. Effects of disruption are assessed in Experiment 2 by comparing two conditions: no disruption and disruption only. In the no-disruption condition, which is typical of past experiments on temporal judgment, subjects remain in the same room during the interval separating D_1 and D_2.

Method

Environmental contexts and materials. The same two rooms were used for the duration contexts as in Experiment 1. Judgments were made in a third context (C_3). Selected to be quite different from both C_1 and C_2, C_3 was a medium-sized room with one door, a linoleum floor, a chalkboard, a U-shaped table, many chairs, and several large windows. It was located one floor below C_1 and C_2. The experimenter in C_3, a male student, was a different person from those in C_1 and C_2. All of the materials were the same as in Experiment 1.

Subjects and design. Subjects were 96 volunteers. Data of an additional 3 subjects were discarded because of their failure to follow instructions. As before, subjects participated in groups. Each group was assigned to one of three conditions: no disruption, disruption only, and disruption and change. A total of 16 subjects in each unchanged-context condition were exposed to each variation of D_1–D_2 contexts (C_1–C_1 or C_2–C_2), and 16 subjects in the disruption-and-change condition were exposed to each variation of D_1–D_2 order of contexts (C_1–C_2 or C_2–C_1). Approximately half of the subjects in each group and exactly half of those in each condition received each duration-judgment form. The variations in condition were run in a counterbalanced order over a 2-day period; at least four groups of subjects contributed data to each condition. Three random orders of assignment of occupations to the two durations were used, and an approximately equal number of subjects in each condition received each version.

Procedure. Subjects assembled for the experiment in C_3. The experimenter led them to the appropriate room for D_1. The instructions and procedure were the same as in Experiment 1 in the disruption-only and disruption-and-change conditions. In the no-disruption condition, subjects were asked simply to wait until the next part of the experiment was set up; they sat quietly in the same room for 120 sec in between D_1 and D_2. At the end of D_2, all subjects returned to C_3 where they made the judgments.

Table 2
Mean D_1/D_2 Ratio in Each Combination of Environmental Context and Judged Duration in Experiment 2

Environmental-context condition	Judged duration D_1	Judged duration D_2	M
No disruption	1.23	1.15	1.19
Disruption only	1.15	.99	1.07
Disruption and change	.97	.95	.96
M	1.12	1.03	1.07

Note. D_1 = first duration; D_2 = second duration.

Results and Discussion

Duration judgment. Each duration judgment was expressed as a D_1/D_2 ratio, and the data were subjected to a 3 × 2 (Environmental Context × Judged Duration) ANOVA and planned t tests. The results, which are shown in Table 2, reveal a major finding: Mean D_1/D_2 ratio is affected by environmental-context condition, $F(2, 90) = 10.6$, $p < .001$, $MS_e = .04$. The means in the no-disruption and the disruption-only conditions are different, $t(62) = 2.31$, $p < .05$, $SE = .05$, as are the means in the disruption-only and the disruption-and-change conditions, $t(62) = 2.12$, $p < .05$, $SE = .05$. The mean ratio of 1.19 in the no-disruption condition reveals a substantial positive time-order error, $t(31) = 5.45$, $p < .001$, $SE = .04$. The mean ratio of 1.07 in the disruption-only condition indicates only a slight positive time-order error, $t(31) = 1.77$, $p = .09$, $SE = .04$. The mean ratio of .96 in the disruption-and-change condition shows no time-order error, $t(31) = 1.20$, $SE = .04$. Although the main effect of judged duration was not significant in Experiment 1, it is here, $F(1, 90) = 4.74$, $p < .05$, $MS_e = .04$. The mean D_1/D_2 ratio is greater for subjects judging D_1 (1.12) than for those judging D_2 (1.03). An explanation of this subsidiary finding is proposed in the Appendix.

List-discrimination judgments. Figure 3 shows mean $P(J_1)$ given recognition in each condition. The overall trends closely replicate those of Experiment 1. Mean $P(J_1)$ is considerably greater for D_1 than D_2 items,

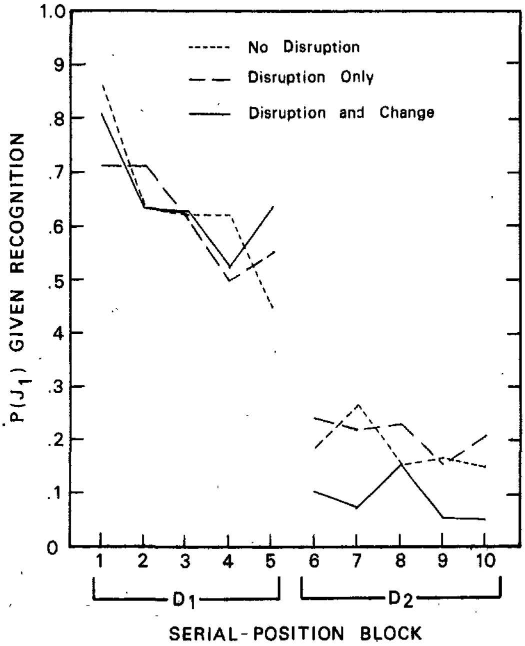

Figure 3. Mean $P(J_1)$ given recognition—probability (P) that an item was judged (J) to have occurred within D_1 given recognition—as a function of serial-position block in each environmental-context condition in Experiment 2. (D_1 = first duration; D_2 = second duration.)

$F(1, 93) = 326$, $p < .001$, $MS_e = .16$. There is a decreasing trend across D_1, $F(1, 93) = 55.6$, $p < .001$, $MS_e = .06$, as well as an interaction of D_1 and D_2 trends, $F(1, 93) = 26.4$, $p < .001$, $MS_e = .04$.

Consider now the between-condition effects and interactions. There is an effect of changing the context prevailing during D_2. The mean $P(J_1)$ on D_2 items is less in the disruption-and-change condition than in either the disruption-only or the no-disruption condition, both $F(1, 62) > 9.19$, $p < .001$, $MS_e < .12$. The difference between D_2 means in the disruption-and-change and the disruption-only conditions replicates the marginally significant effect found in Experiment 1. Changing the context makes it somewhat less likely that D_2 items will be incorrectly assigned to D_1. However, there are no significant differences between the no-disruption and the disruption-only conditions, all $F(1, 62) < 2.24$. Although the disruption was powerful enough to affect duration judgment, it did not improve list discrimination.

Serial-position judgments. Figure 4 shows the mean position judgment of each kind in the three conditions. Given a correct list judgment of D_1 items ($J_1|D_1$), the overall slope of .28 is greater than zero, $t(3) = 2.62$, $p < .05$; given a correct judgment of D_2 items ($J_2|D_2$), the slope of .20 is also greater than zero, $t(3) = 2.51$, $p < .05$. However, given an incorrect list judgment of D_1 items ($J_2|D_1$), the slope of .05 is not significantly different from zero $t(3) = 1.44$; and given an incorrect judgment of D_2 items ($J_1|D_2$), the slope of .03 is also not significantly different from zero, $t(3) = .23$. These findings, which essentially replicate these of Experiment 1, show that serial-position information was available, but only if the list was correctly identified.

There are no significant differences between corresponding slopes in the three conditions, all $t(6) < 1.97$, and the difference between the $J_1|D_1$ and $J_2|D_2$ slopes is not significant in any condition, all $t(6) < 1.12$. Fur-

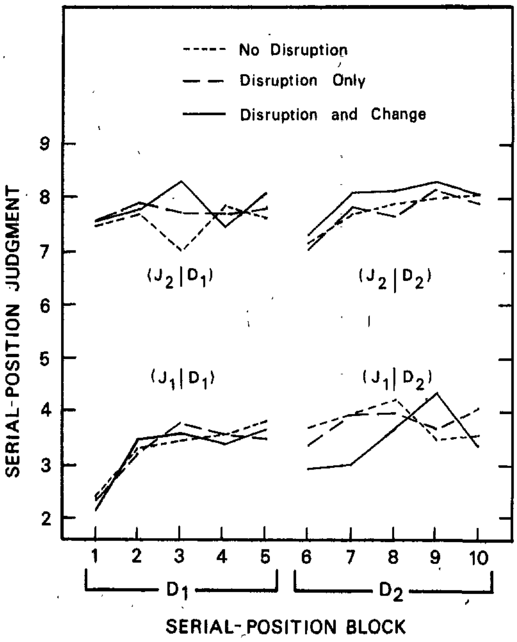

Figure 4. Mean serial-position judgment of each kind as a function of serial-position block in each environmental-context condition in Experiment 2. ($J_1|D_1$ = correct list judgment for first-duration item. $J_2|D_2$ = correct list judgment for second-duration item. $J_2|D_1$ = incorrect list judgment for first-duration item. $J_1|D_2$ = incorrect list judgment for second-duration item.)

thermore, given a correct list judgment, the D_1/D_2 slope ratio is 1.39 in the no-disruption, 1.28 in the disruption-only, and 1.69 in the disruption-and-change condition. As before, the pattern of ratios is clearly different from that of the D_1/D_2 duration-judgment ratios.

Summary. If there is no disruption of environmental context, a substantial positive time-order error in duration judgment is found; if there is a disruption, the error is reduced; and if there also is a change, it is eliminated. List discrimination, on the other hand, is affected by a change in context prevailing during D_2, but not by a disruption of context during the interval separating D_1 and D_2. Neither disruption nor change of context affected position judgments.

Experiment 3

Experiments 1 and 2 show that comparative-duration judgment is affected both by a disruption in environmental context during the interval between D_1 and D_2 and by a change prevailing during D_2. As noted earlier, Block and Reed (1978, Experiment 2) found that duration judgment is also affected by process-context changes during a duration. The process-context manipulation was accomplished by instructing subjects to process information at a single level (semantic or structural) during one duration—an unmixed-processing task—and to process information at different levels (alternately semantic and structural) during another duration—a mixed-processing task. Half of their subjects performed mixed processing during D_1 (mixed–unmixed condition) and half during D_2 (unmixed–mixed condition). If the duration judgment is expressed as a ratio of the remembered duration of the mixed-processing task to that of the unmixed-processing task, the ratio is greater than 1.00 in the mixed–unmixed condition but is not different from 1.00 in the unmixed–mixed condition. Overall, the mixed duration is judged to be about 12% longer than the unmixed duration. If each judgment is expressed as a D_1/D_2 ratio, there is a positive time-order error in the mixed–unmixed condition but not in the unmixed–mixed condition.[2] Overall, D_1 is judged to be about 15% longer than D_2. Block and Reed suggested that an effect of mixed processing is added to a time-order error in the mixed–unmixed condition but is subtracted in the unmixed–mixed condition. The present experiment addresses the question of whether or not different kinds of contextual changes are additive (or subtractive) in their effect on remembered duration, as implied by Block and Reed. It does so by orthogonally combining the environmental-context manipulation of the present experiments and the same kind of process-context manipulation as that of Block and Reed.

Method

Environmental contexts. Contexts C_1 and C_2 were used for the duration contexts, and C_3 was used for the judgment context.

Materials. A total of 17 words was generated in each of four categories: small-size living (e.g., mouse), large-size living (e.g., whale), small-size nonliving (e.g., bullet), and large-size nonliving (e.g., piano). Four words from each were randomly assigned to each of four type styles: uppercase regular, uppercase italic, lowercase regular, and lowercase italic. One word from each set of four was assigned to each of four semantic (deep) descriptions, which corresponded to the categories (e.g., large-size living), and also assigned to each of four structural (shallow) descriptions, which corresponded to the type styles (e.g., capital italic type). Thus, only one word of each set of four correctly matched the description of it. For use in C_1, each word was typed (using either an IBM Prestige Elite or an IBM Light Italic style) and mounted in a slide frame. One of the two descriptions, either shallow or deep, was typed (using uppercase IBM Orator type) above the word. For use in C_2, each word was printed (using either Chartpak 36-point American Typewriter Medium letters or Bookman Bold Italic and Commercial Script letters) on an index card. One of the two descriptions was printed (using a 10-mm Rapidograph lettering guide) above the word. A total of 32 words was randomly assigned to each of two lists so that there was an equal number with each possible combination of category, type style, and description. Thus, in each list a total of 8 words correctly matched the description (e.g., large-size living for the word *giraffe*); 16 differed from the description in a single feature (e.g., small-size nonliving for the word *ant*); and 8 differed in both features (e.g., small-size living for the word *house*). The remaining word in each category was used only as a distractor item on the subsequent memory test.

Mixed-processing lists were constructed by using both

[2] This analysis was not reported by Block and Reed (1978). The mean ratio of 1.24 in the mixed–unmixed condition is significantly greater than 1.00, $t(31) = 4.50$, $p < .001$, $SE = .05$, but the mean of 1.05 in the unmixed–mixed condition is not, $t(31) = 1.25$, $SE = .04$. The two means are significantly different, $F(1, 56) = 8.12$, $p < .01$, $MS_e = .07$.

16 shallow and 16 deep descriptions; the two kinds of description occurred in a random order. Unmixed-processing lists were constructed by using either 32 shallow or 32 deep descriptions; the particular shallow or deep descriptions occurred in a random order.

As before, two versions of a duration-judgment form were used; the instructions were slightly modified. The memory-judgment form was similar to that in Experiments 1 and 2. One difference was that it contained 36 words with 4 words randomly selected from each quarter of each list and 4 distractor words. Each word was typed in the style to which it had been assigned. Another difference was that the position-judgment scale was divided into eight, rather than 10, categories; thus, possible responses ranged from 1 (start of first series of words) to 8 (end of second series of words).

Subjects and design. Subjects were 192 volunteers. Data from an additional 6 were discarded because of their failure to follow instructions. Subjects participated in groups. Each group was assigned to one of eight conditions; a total of 24 subjects served in each. The conditions were an orthogonal combination of environmental context (disruption-only or disruption-and-change), process-context order (mixed–unmixed or unmixed–mixed), and unmixed-processing type (shallow or deep). The eight conditions were run in a counterbalanced order on both of two days. Between days, assignment of lists to mixed- and unmixed-processing conditions was rotated by substituting the alternate descriptions for half of the words.

Procedure. Subjects assembled in C_3, and they were led to the appropriate room for D_1. The instructions and general procedure were similar to that in Experiments 1 and 2. Subjects were told that the task involved "judging whether or not a certain description matches a word." They were fully informed of the four possible shallow and four possible deep descriptions and were shown examples of one shallow- and one deep-description slide or card. They were told that "when the description concerns the category of the word," they "need not be concerned with [its] typestyle," and vice versa. Subjects were told to write quickly each decision, "yes" or "no," on a separate page of the notepad. The instructions lasted 90 sec. Items were presented at a 5-sec rate; thus, each duration was 160 sec. Between D_1 and D_2, subjects waited in the hallway for 80 sec. At the end of D_2, they returned to C_3 to make the judgments.

Results and Discussion

Duration judgment. Each duration judgment was expressed initially as a D_1/D_2 ratio, and the data were subjected to a 2 × 2 × 2 × 2 (Environmental Context × Process Context × Unmixed Task × Judged Duration) ANOVA. The ANOVA reveals an important interaction between environmental and process context, $F(1, 184) = 5.55, p < .05, MS_e = .06$. Planned t tests clarify this finding: The mean ratio in the disruption-only, mixed–unmixed condition shows a positive time-order error, $t(47) = 4.01, p < .001, SE = .03$, but none of the other conditions shows a time-order error, all $t(47) < 1.10$. The only other significant effect is a main effect of process context, $F(1, 184) = 4.94, p < .05, MS_e = .06$.

Table 3
Mean D_1/D_2 Ratio in Each Combination of Environmental Context, Process Context, and Judged Duration

Environmental- and process-context condition	Judged duration		
	D_1	D_2	M
No disruption			
Mixed–unmixed	1.30	1.18	1.24
Unmixed–mixed	1.13	.97	1.05
M	1.22	1.08	1.15
Disruption only			
Mixed–unmixed	1.10	1.15	1.13
Unmixed–mixed	.93	1.00	.97
M	1.02	1.08	1.05
Disruption and change			
Mixed–unmixed	1.05	1.02	1.04
Unmixed–mixed	1.08	1.00	1.04
M	1.07	1.01	1.04

Note. The no-disruption condition is that of Block and Reed (1978, Experiment 2); the disruption-only and disruption-and-change conditions are those of the present Experiment 3. D_1 = first duration; D_2 = second duration.

Mean D_1/D_2 ratios in the Block and Reed (1978; no-disruption) and the present (disruption-only and disruption-and-change) conditions are summarized in Table 3. Note that process context does not affect the D_1/D_2 ratio if environmental context is changed, only if it is unchanged. Note also that environmental context affects the D_1/D_2 ratio if process context is changed during D_1 (i.e., mixed–unmixed condition), but not if it is changed during D_2 (i.e., unmixed–mixed condition). The effect of environmental context in the former is strikingly similar to that in the present Experiments 1 and 2. However, the apparent lack of an effect of environmental context in the latter is potentially misleading. The D_1/D_2 ratios should be less than 1.00 if environmental context is not important but process context is. Because the ratios are not less than 1.00, environmental context must be important even if process context is changed during D_2.

In order to clarify these findings, each duration judgment was expressed as a ratio of the apparent duration of the mixed-processing task to that of the unmixed-processing task (D_m/D_u ratio). A 2 × 2 × 2 × 2 ANOVA reveals that environmental context affects the ratio, $F(1, 184) = 4.52$, $p < .05$, $MS_e = .09$. The mean D_m/D_u ratio is greater if environmental context is only disrupted than if it is changed. The mean ratio of 1.13 in the disruption-only condition is greater than 1.00, $t(95) = 3.76$, $p < .001$, $SE = .03$, but the mean ratio of 1.03 in the disruption-and-change condition is not, $t(95) = 1.22$, $SE = .03$.

In summary, the major finding is that of an interaction between environmental and process context, which suggests that effects of the two manipulations are not additive. This interaction and the other findings can be explained by a clarification of the contextual-change hypothesis that is described in the General Discussion section.

List-discrimination judgments. Figure 5 shows mean $P(J_1)$ given recognition in the two environmental-context conditions. Consider first the overall trends, which quite closely replicate the findings of Experiments 1 and 2. Mean $P(J_1)$ is greater for D_1 words than for D_2 words, $F(1, 184) = 110$, $p < .001$, $MS_e = .11$. Although performance is not as good as in Experiments 1 and 2, subjects nevertheless showed considerable list-discrimination ability. The interaction of D_1 and D_2 trends is marginal, $F(1, 184) = 4.89$, $p = .03$, $MS_e = .08$, replicating that in Experiments 1 and 2.

Consider now the between-condition effects and interactions. A major finding is that there is an effect of environmental context, as revealed by an interaction on the D_1–D_2 mean comparison, $F(1, 184) = 7.20$, $p < .01$, $MS_e = .14$. Mean $P(J_1)$ on D_1 items is not significantly different between the two conditions, $F(1, 184) = 1.15$, $MS_e = .20$; however, mean $P(J_1)$ on D_2 items is marginally lower in the disruption-and-change than in the disruption-only condition, $F(1, 184) = 4.42$, $p = .03$, $MS_e = .16$. This finding supports the conclusions of Experiments 1 and 2: Changing the environmental context makes it somewhat less likely that D_2 items will be incorrectly assigned to D_1. It is also notable

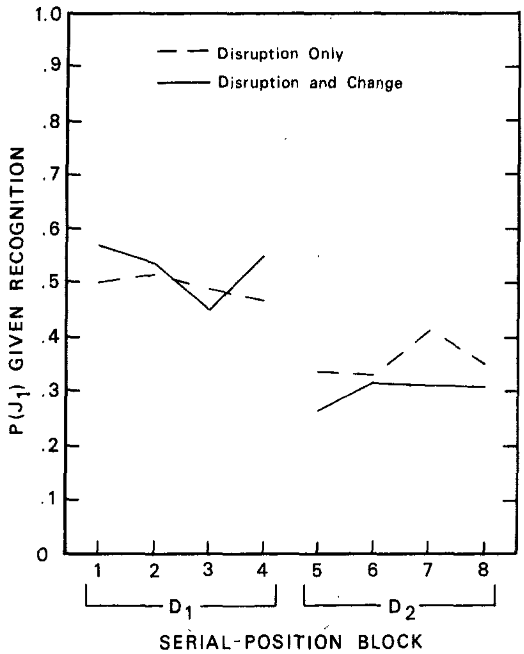

Figure 5. Mean $P(J_1)$ given recognition—probability (P) that an item was judged (J) to have occurred within D_1 given recognition—as a function of serial-position block in each environmental-context condition in Experiment 3. (D_1 = first duration; D_2 = second duration.)

that there is nothing distinctive about performance in the disruption-only, mixed–unmixed condition, the only condition to show a time-order error in duration judgment.

Although there are no other significant effects or interactions involving environmental context, there are some involving process-context order and unmixed-task condition. There is an effect of unmixed task on mean $P(J_1)$ for D_1 words, as well as an interaction between process context and unmixed task on both mean $P(J_1)$ for D_2 words and the comparison of D_1 and D_2 means, all $F(1, 160) > 9.03$, $p < .005$. All of these findings can be attributed to a tendency correctly to assign D_1 words to D_1 more frequently if the words had received shallow processing (M = .61) instead of mixed or deep processing (Ms = .49 and .45, respectively) and a tendency incorrectly to assign D_2 words to D_1 more frequently if the words had received shallow processing (M = .45) instead of mixed or deep processing (Ms = .31 and .24, respectively). Thus, it seems that shallowly pro-

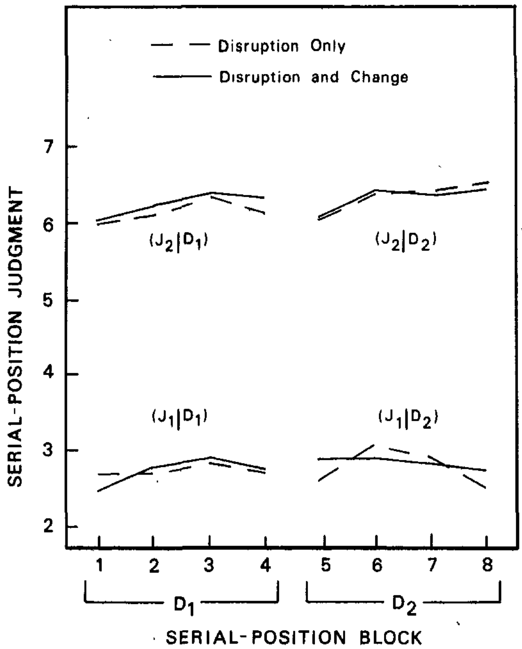

Figure 6. Mean serial-position judgment of each kind as a function of serial-position block in each environmental-context condition in Experiment 3. ($J_1|D_1$ = correct list judgment for first-duration item. $J_2|D_2$ = correct list judgment for second-duration item. $J_2|D_1$ = incorrect list judgment for first-duration item. $J_1|D_2$ = incorrect list judgment for second-duration item.)

cessed items tend to be remembered as relatively less recent than deeply processed items (cf. Block & Reed, 1978). There are also typical large effects on recognition (see Appendix).

Serial-position judgments. Judgments were pooled as before, and the mean judgment of each kind is shown in Figure 6. Given a correct list judgment of D_1 items ($J_1|D_1$), the overall slope is .05; and given a correct list judgment of D_2 items ($J_2|D_2$), the slope is .12. Given an incorrect judgment of D_1 items ($J_2|D_1$), the overall slope is .09; and given an incorrect judgment of D_2 items ($J_1|D_2$), the slope is −.05. None of these slopes is significantly greater than zero, all $t(2) < 1.92$, indicating that little or no serial-position information was available at the time of the test. Analysis of the slopes in the four Environmental-Context × Process-Context conditions reveals no significant difference between any of the corresponding slopes, all $t(4) < 1.21$. There is nothing distinctive about the slopes in the disruption-only, mixed–unmixed condition.

General Discussion

Considered together, the results of the three experiments reported here are fairly consistent. Comparative-duration judgment is affected by environmental- and process-context changes. If there is no disruption of environmental context, D_1 is remembered as being longer than D_2 (a positive time-order error); if the context is disrupted during the interval separating D_1 and D_2, the time-order error is reduced; if the environmental context during D_2 is changed from that during D_1, the error is eliminated. Overall, a duration containing process-context changes is remembered as being longer than one containing no such changes. However, the finding of an interaction between environmental- and process-context change suggests that the two kinds of contextual change do not produce additive effects.

In contrast, list discrimination is slightly improved if a changed context prevails during D_2, but not if context is only disrupted. List discrimination is affected by the process-context manipulation only to the extent that items were processed at different levels. Serial-position judgment apparently is affected by neither the environmental- nor the process-context manipulation used here.

Theoretical Implications

Duration-judgment processes. The finding of effects of environmental- and process-context changes supports a contextual-change hypothesis on remembered duration. The typical finding of a positive time-order error in duration judgment suggests that contextual encoding changes more if a novel context prevails during a time period, as it ordinarily does during D_1. The finding here that the time-order error is eliminated if a novel context also prevails during D_2 supports this assumption. Because the time-order error is also reduced slightly if there is a mere disruption of context during the interval separating D_1 and D_2, it appears that contextual encoding during D_2 may be affected by such a disruption. The disruption may cause either the renewed encoding of familiar contextual elements or the encoding of novel elements during D_2. A contextual-change hypothesis is also supported by the finding that remembered duration is lengthened if there are process-context changes during a time period.

However, the interaction between environmental and process context in the data of Experiment 3 and of Block and Reed's (1978) Experiment 2 suggests that changes in environmental and process context are not additive. D_1 is remembered as being longer than D_2 if process context changes occur during D_1 and environmental context is unchanged; otherwise, there is no time-order error. One possible explanation is that relatively greater changes in the encoding of both process and environmental context during D_1 as compared with D_2 are necessary to produce the time-order error. This explanation, however, cannot easily explain the finding of a positive time-order error in experiments in which neither process nor environmental context was varied (Block, 1978). Another possible explanation is that the time-order error is primarily a result of the relative amount of change in the encoding of environmental context during each duration. Other kinds of contextual elements, such as those encoding process context, may be salient only to the degree that there is no change in encoding of environmental context during D_2. Thus, remembered duration may be a result of a process of retrieving salient contextual elements associated with a time period, followed by a process of estimating the amount of change in the retrieved elements. The overall process may be similar to that which is assumed to underlie the use of an availability heuristic in frequency judgment (Tversky & Kahneman, 1973).

List-discrimination processes. List-discrimination judgments apparently are mediated by somewhat different aspects of memory from duration judgment, because the two kinds of judgment were affected in different ways by context manipulations in Experiments 2 and 3. List discrimination is apparently based on a process of retrieving contextual elements associated with an event, followed by a process of deciding whether the retrieved elements are more likely to have been encoded during one time period or the other. The latter process is affected by the overall similarity of the set of contextual elements associated with D_1 to the set associated with D_2, and inference processes are implicated (Hintzman et al., 1973; Underwood, 1977). List-discrimination performance may serve as an index of the amount of contextual similarity between two time periods, but not necessarily as an index of change in contextual encoding during a time period.

Position-judgment processes. The serial-position data are somewhat puzzling. A contextual-association hypothesis (see Hintzman & Block, 1971; Hintzman et al., 1973) proposes that the position-judgment slope reflects the amount of contextual change during a time period, just as duration judgment apparently does. However, none of the present experiments reveals any effect on position judgment of the kinds of manipulations used here. Serial-position judgment might depend on a process of retrieving contextual elements associated with an event, followed by a process of inferring the most likely relative time of encoding of the elements. Neither environmental- nor process-context elements seem to be involved, although the reason is unclear. Future experiments should gather more data relevant to the relationship between serial-position judgment and other kinds of temporal judgment.

A Theoretical Synthesis

The three kinds of temporal judgment explored here are apparently based on somewhat different aspects of memory, because the context manipulations had different effects on each kind of judgment. The explanations offered have focused mainly on retrieval and judgment processes. Consider now a possible way to explain the findings in terms of transient changes in encoded contextual information that influence each kind of temporal judgment (cf. Hintzman et al., 1973). Assume that the set of contextual elements prevailing at any moment consists of three subsets: A, B, and C. Elements in Subset A encode environmental-context information. If the environmental context is not disrupted, the rate of change of elements in A is a negatively accelerated function of time in the context, as fewer novel elements are encoded. If environmental context is disrupted during the interval separating D_1 and D_2, there are discontinuities in the function: increased rates of change attributable to the encoding of the disruption context and the renewed encoding of familiar elements during D_2. If the environmental context during D_2 differs from that during D_1, there is an additional discontinuity in the function: an increased rate of change attributable to the encoding of novel elements during D_2. Elements in Subset B encode process-context information. Their average rate of change is greater during a mixed-processing task than during an unmixed-processing task. The rate also increases at the start of the interval separating D_1 and D_2, as the D_1 task ends and the subject waits for D_2; depending on the D_2 task, some or all of the original (D_1) elements might be reinstated at the start of D_2. Elements in Subset C encode within-duration information. They change in a regular way during each task, but the original elements are reinstated at the start of the D_2 task. The present experiments do not reveal the specific kind of elements involved.

All of the major findings may now be explained by summarizing and clarifying the previously mentioned differences in retrieval and judgment processes. Comparative-duration judgment may involve an assessment of the amount of change in some or all of the three subsets during D_1 relative to that during D_2. The amount of change in Subset A is salient in making this kind of judgment. Elements in Subset B may become salient to the degree that there is little or no change in Subset A during D_2. List-discrimination judgment may involve an assessment of the differences between D_1 and D_2 elements in Subset A. (Plotting the curves in Figures 1, 3, and 5 on a true temporal scale reveals how Subset A may account for the major trends observed.) To the extent that the processing task performed during D_2 is different from that during D_1, elements in Subset B may also be involved. (In the present Experiment 3, however, both durations contained some of the same level of processing.) Serial-position judgment apparently involves an assessment of differences among elements in Subset C. It is unclear why Subsets A and B are apparently not considered.

The experiments reported here have been a fruitful exploration of contextual factors involved in temporal memory judgments. Future studies along similar lines might help to restore cognitive context to an important role in explanations of learning, memory, and cognitive processes.

References

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

Block, R. A. Time and consciousness. In G. Underwood & R. Stevens (Eds.), *Aspects of consciousness* (Vol. 1). London: Academic Press, 1979.

Block, R. A., George, E. J., & Reed, M. A. A watched pot sometimes boils: A study of duration experience. *Acta Psychologica*, 1980, *46*, 81–94.

Block, R. A., & Reed, M. A. Remembered duration: Evidence for a contextual-change hypothesis. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, *4*, 656–665.

Crowder, R. G. *Principles of learning and memory*. Hillsdale, N.J.: Erlbaum, 1976.

Hewitt, K. *Contextual aspects of memory in normal and pathological states*. Unpublished master's thesis, Cambridge University, 1979.

Hintzman, D. L., & Block, R. A. Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology*, 1971, *88*, 297–306.

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

McGeoch, J. A. *The psychology of human learning: An introduction*. New York: Longmans, Green, 1942.

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

Rosenthal, R. Combining results of independent studies. *Psychological Bulletin*, 1978, *85*, 185–193.

Smith, S. M. Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory*, 1979, *5*, 460–471.

Smith, S. M., Glenberg, A., & Bjork, R. A. Environmental context and human memory. *Memory & Cognition*, 1978, *6*, 342–353.

Strand, B. Z. Change of context and retroactive inhibition. *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 202–206.

Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 1973, *5*, 207–232.

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

Zechmeister, E. B., & Nyberg, S. E. *Human memory: An introduction to research and theory*. Monterey, Calif.: Brooks/Cole, 1982.

Appendix

Subsidiary Data Analyses

Duration judgment. In all three experiments, each duration judgment was measured in millimeters prior to converting it to a D_1/D_2 ratio and collapsing over judged-duration condition. The subsidiary analyses reported here offer additional insight into processes involved in comparative duration judgment and also reinforce the previous conclusions based on the D_1/D_2 ratio.

The overall mean judgment is 52.7, 52.9, and 51.4 mm in Experiments 1, 2, and 3. The mean is significantly greater than 50.0 mm in both Experiment 1, $t(111) = 2.60, p < .05, SE = 1.1$, and Experiment 2, $t(95) = 2.47, p < .05, SE = 1.2$, but not in Experiment 3, $t(191) = 1.49, p > .05, SE = .09$. Considering only subjects for whom D_1 was the judged duration, the mean judgment is 55.3, 55.3, and 52.6 mm in Experiments 1, 2, and 3. The mean is significantly greater than 50.0 mm in both Experiment 1, $t(55) = 3.79$, $p < .001$, $SE = 1.3$, and Experiment 2, $t(47) = 2.98$, $p < .01$, $SE = 1.8$, and marginally greater in Experiment 3, $t(95) = 1.89, p = .06, SE = 1.4$. Considering only subjects for whom D_2 was the judged duration, the mean judgment is 50.0, 50.5, and 50.2 mm in the three experiments. The mean is not different from 50.0 mm in any of them (all $ts < .36$, $SE < 1.7$). The difference between the mean for subjects judging D_1 and that for subjects judging D_2 is significant in both Experiment 1, $F(1, 108) = 6.94, p < .05, MS_e = 112$, and Experiment 2, $F(1, 90) = 4.92, p < .05, MS_e = 110$, and marginal in Experiment 3, $F(1, 176) = 3.14$, $p = .08, MS_e = 160$. To my knowledge, this is the first time that such effects have been reported when a method of comparative-duration judgment was used. An understanding of the source of these effects might help to explain the typical finding of a positive time-order error.

To ascertain the reliability of these data, they were combined with all data from appropriately designed experiments conducted in my laboratory (i.e., Block, 1978, Experiments 1 and 2; Block & Reed, 1978, Experiments 1 and 2), with all other variables counterbalanced. The overall mean judgment of 52.6 mm is significantly greater than 50.0 mm ($Z = 4.55$, $p < .001$, one-tailed), using the method of adding weighted Zs (Rosenthal, 1978). Across all subjects judging D_1, the mean of 55.8 mm is greater than 50.0 mm ($Z = 5.55$, $p < .001$, one-tailed). Across all subjects judging D_2, the mean of 49.3 mm is not significantly different from 50.0 mm ($Z = -.78$). In short, subjects make a longer comparative duration judgment if they judge D_1 in comparison with D_2, but not if they judge D_2 in comparison with D_1. This finding, at least to some extent, reflects the positive time-order error. It may be attributed to either a response bias, a genuine effect on remembered duration, or both. The interpretation that is most consistent with the evidence reported here is that it is a genuine effect. A possible explanation is that there is a greater focusing of attention on, or salience of, contextual elements that are associated with the judged duration. Whatever the best explanation is, the finding should caution researchers to ensure appropriate counterbalancing on the judgment task, as well as to avoid other theoretical and statistical pitfalls (see, e.g., Ornstein's, 1969, treatment of his data).

Recognition judgments. Like other kinds of temporal memory judgment, recognition-memory judgment can be considered to depend on available contextual information. Specifically, the retrieval process used in making a recognition judgment might assess whether or not there is any available association between an event and the context defining the decision, such as the general experimental context. The present experiments, however, were not designed to test this contextual hypothesis on recognition judgment, but rather to produce high recognition rates. Thus, the data reported here are subsidiary to the primary analyses.

In Experiment 1, the overall mean probability that a presented occupation was correctly recognized, the hit rate, is .98, and the overall mean probability that a nonpresented occupation was incorrectly recognized, the false-alarm rate, is .09. Each subject's recognition performance was measured at each serial-position block by using a standard correction for guessing, which is the hit rate at the serial-position block minus the false-alarm rate. There are no significant effects or interactions. It is possible, of course, that the extremely accurate recognition performance obscured the finding of any significant effects.

In Experiment 2, the overall mean hit rate is .98, and the overall mean false-alarm rate is .05. As in Experiment 1, there are no significant effects or interactions, but there is the possibility of a ceiling effect.

In Experiment 3, the overall mean hit rate is .73, and the overall mean false-alarm rate is .24. Analysis of corrected recognition performance reveals that the overall D_2 mean of .50 is marginally greater than the D_1 mean of .47, $F(1, 184) = 6.16$, $p = .014$, $MS_e = .06$, replicating a similar finding of Block and Reed (1978). There is also an increasing trend across D_2, $F(1, 184) = 6.98$, $p < .01$, $MS_e = .03$. Both findings reflect a recency effect.

Between conditions, there is no significant main effect of either environmental context, all $F(1, 184) < 2.10$, or process context, all $F(1, 184) < 1.98$. As expected, however, there are main effects of unmixed task and interactions of process context and unmixed task on both mean D_1 and D_2 performance, all $F(1, 184) > 14.0$, $p < .001$. Simply stated, performance is high on items from an unmixed-deep duration (overall $M = .68$), moderate on items from a mixed-processing duration (overall $M = .52$), and low on items from an unmixed-shallow duration (overall $M = .31$). These findings replicate those of Block and Reed (1978); the difference between deep- and shallow-processing tasks has also been found by many other investigators.

List-discrimination judgments. As Hintzman et al. (1973) found, subjects do not use position-judgment categories proportionally. In Experiment 1, the overall probabilities of use of Judgment Categories 1 through 10 for correctly recognized (old) items are .04, .07, .10, .11, .08, .11, .14, .15, .13, and .08. In Experiment 2, the probabilities are .04, .07, .09, .12, .07, .10, .16, .16, .13, and .05. The probabilities for Categories 1 through 8 in Experiment 3 are .05, .12, .14, .10, .15, .19, .18, and .07. In all three experiments, these data show a bias to use D_1 categories less frequently than D_2 categories. There is no significant between-condition interaction in any of them (all $Fs < 1$). Data from all three also reveal a bias to use D_1 categories more frequently than D_2 categories in judgments of incorrectly recognized (new) items. In Experiment 1, these overall probabilities for Categories 1 through 10 are .07, .07, .21, .18, .11, .07, .07, .10, .05, and .08. In Experiment 2, they are .05, .10, .33, .14, .05, .05, .05, .19, .05, and .00. In Experiment 3, for Categories 1 through 8, they are .12, .23, .29, .08, .09, .13, .03, and .03.

Received October 8, 1981
Revision received April 12, 1982 ■