Perception & Psychophysics
1985, 38 (2), 115-124

Time perception and attention: The effects of prospective versus retrospective paradigms and task demands on perceived duration

SCOTT W. BROWN
University of Maine, Orono, Maine

This research was designed to compare time judgments obtained under prospective conditions (in which subjects are instructed to attend to time) and retrospective conditions (in which subjects are unaware that they will be required to judge time). In Experiment 1, subjects prospectively or retrospectively judged the duration of intervals spent performing a perceptual-motor task at different levels of difficulty. The results showed that subjects tested under both research paradigms tended to give increasingly shorter and/or more inaccurate time judgments with increases in nontemporal task demands. Experiment 2 was designed to test the effects of attentional deployment on perceived time by comparing prospective and retrospective judgments under control, selective attention, and divided attention conditions. Both types of time judgments became increasingly inaccurate as attention was more broadly deployed. The results of these experiments are consistent with an attentional allocation model, and they suggest that nontemporal task demands disrupt or interfere with timing in both prospective and retrospective situations.

Many writers have noted that a major factor affecting the experience of duration is the degree to which attention is directed to the flow of time itself (e.g., Fraisse, 1984; Frankenhaeuser, 1959; James, 1890/1950; Sturt, 1925). This heightened awareness of the passage of time has been termed "the experience of time-in-passing" (Hicks, Miller, Gaes, & Bierman, 1977). Situations involving a heightened temporal awareness, such as boredom, impatience, and anticipation, often seem to produce an apparent lengthening (or slowing down) of external time. The classic example of this effect is the "watched-pot phenomenon" (Fraisse, 1963), where time seems to drag slowly by. In contrast, when one is engaged in some absorbing activity, temporal awareness becomes relatively minimized and perceived duration becomes shortened. Various laboratory studies designed to enhance or minimize temporal awareness tend to confirm these subjective impressions (e.g., Cahoon & Edmonds, 1980; Curton & Lordahl, 1974; McKay, 1977).

One important factor in evaluating the experimental literature is the methodological distinction between prospective and retrospective research paradigms. In the prospective paradigm, subjects are explicitly told in advance that they will be required to judge the duration of an interval. This procedure presumably motivates subjects to monitor the time going by and to attend to any available temporal cues (Doob, 1971). In contrast, subjects tested under the retrospective paradigm are not given any prior warning about time judgments at the start of the interval. These subjects are unexpectedly asked to judge the duration of the interval after it has already passed by. Subjects tested under retrospective conditions are presumed to process temporal information in a more incidental and unreliable fashion.

Gilliland, Hofeld, and Eckstrand (1946) first raised the question as to whether prospective and retrospective conditions would have different effects on duration judgments. Unfortunately, research with the two paradigms has tended to follow separate lines of investigation. By far, most of the research literature is prospective in nature. Much of this work concerns the effects of nontemporal processing on prospective timing. Several theorists (Hicks, Miller, & Kinsbourne, 1976; Thomas & Brown, 1974; Thomas & Cantor, 1978; Thomas & Weaver, 1975) have applied an attentional allocation or distraction model to prospective situations during which subjects were also required to perform some attention-demanding task. According to the Thomas model, subjects are faced with a dual task and share attention between temporal and nontemporal processing. As nontemporal task demands increase, less attentional capacity is allocated to temporal processing, and duration judgments become more unreliable. Hicks et al. (1976) expressed a similar idea by proposing that nontemporal demands serve to distract one from processing temporal cues; as a result of fewer cues being stored, time judgments become shorter. The two theories differ in certain details: The Thomas model, originally developed for intervals on a millisecond time scale, predicts greater un-

This article is based on the author's doctoral dissertation from the University of Maine. A shortened version of Experiment 1 was presented at the annual meeting of the Eastern Psychological Association, Boston, March 21-24, 1985. I gratefully acknowledge the invaluable assistance of D. Alan Stubbs, who contributed in many ways to all phases of this project.

Requests for reprints should be sent to Scott W. Brown, Department of Psychology, University of Maine, Orono, ME 04469.

reliability in time judgments, whereas Hicks et al. focus on longer intervals and predict greater underestimation. However, both theories agree that nontemporal processing demands disrupt or interfere with timing by diverting attention away from the temporal task.

The results of many prospective time studies can be interpreted within this general framework. When time judgments of intervals spent performing some effortful or difficult task are compared against judgments of intervals in which the only task was to keep track of time, involvement in a nontemporal task generally shortens perceived duration (Axel, 1924; Burnside, 1971; Cohen, 1971; DeWolfe & Duncan, 1959; Gulliksen, 1927; Hawkes, 1972; Hawkes & Sherman, 1972; Hicks & Brundige, 1974; Wilsoncroft & Stone, 1975). Further increases in nontemporal processing demands lead to a progressive shortening of perceived time (Allen, 1980; Harton, 1938a, 1938b, 1942; Hicks et al., 1977; Hicks et al., 1976; Smith, 1969; Zakay, Nitzan, & Glicksohn, 1983). An attentional distraction model explains these results in terms of the relative amount of attentional capacity devoted to timekeeping. The more capacity needed for the nontemporal task, the less capacity available for attending to time.

Attention to time-in-passing presumably would not be a factor in retrospective situations. Instead, retrospective judgments would be based upon whatever temporal information had been processed and stored as part of an overall temporal framework in which sequences of events and experiences are organized (e.g., Block, 1979). Consequently, the relatively few retrospective studies that exist are primarily concerned with the role of memory on perceived time (see Ornstein, 1969). However, some retrospective studies can be viewed in terms of attentional demands.

In Ornstein's (1969) research, subjects retrospectively compared intervals involving perceptual-motor tasks (Experiment 5) or complex categorization tasks (Experiment 6) against standard control intervals. The more difficult or complex the task, the longer its perceived duration. A similar outcome was obtained by Underwood (1975), who had subjects judge the relative durations of intervals spent studying and recalling a word list. The retrieval phase was judged longer than the encoding phase. Since numerous studies indicate that retrieval requires more processing capacity (e.g., Martin, 1970; Trumbo & Milone, 1971), this result suggests that task demands lengthen retrospective time judgments.

Other studies, however, have produced contrary results. Vroon (1970) manipulated the amount of information processed by subjects in a tone-discrimination task and found that conditions requiring more processing were judged as shorter than conditions requiring less processing (see also Block & Reed, 1978). The different studies show that the effects of nontemporal processing demands on retrospective timing is unclear. One obstacle in evaluating this literature is a lack of relevant studies.

One promising approach to the study of attention and time would be to compare prospective and retrospective time judgments directly across similar task conditions. Taken separately, prospective and retrospective studies are often difficult to compare because they usually differ considerably with regard to the nontemporal task, the duration, and the nature of the judgment method. Thus, a comparison of the two paradigms within the same experiment may help clarify the role of attentional processes in prospective and retrospective timing. To date, only a handful of studies have been specifically designed to compare the two paradigms. Hicks et al. (1976) had subjects judge the duration of card-sorting tasks during which they had processed varying amounts of information. Prospective time judgments decreased monotonically with increases in processing demands, whereas retrospective judgments did not change significantly. McClain (1983) conducted a paradigm-comparison study in which she varied both the difficulty and the number of word stimuli presented to subjects in a semantic-processing task. Prospective judgments decreased with increases in processing demands, whereas processing demands had no effect on retrospective judgments. The retrospective judgments seemed to be influenced primarily by the number of stimuli presented during the interval; the more words presented, the longer the retrospective judgments.

One of the most ambitious studies of this type was performed by Miller, Hicks, and Willette (1978), who attempted to manipulate task difficulty by having subjects study and recall a word list for varying numbers of trials. The reasoning was that the amount of processing capacity expended for the memorization task would be an inverse function of the number of previous trials. On the study trial for which time estimates were to be obtained, the subjects were told either to actively rehearse the words or to "rest" during the interval. For the nonrehearsal condition, prospective judgments were greater than retrospective judgments, an effect consistent with the idea that prospective subjects were actively attending to time-in-passing. For the rehearsal condition, prospective judgments increased linearly as a function of the number of previous trials. According to Miller et al. (1978), this effect occurred because subjects were increasingly able to direct more processing capacity away from the memorization task and allocate it to the timing task. The judgments of retrospective subjects in the rehearsal condition decreased linearly as a function of the number of previous study trials. This finding was interpreted to mean that retrospective judgments were based on the amount of processing occurring during the interval, so that greater expenditures of processing capacity are associated with longer retrospective judgments and vice versa.

Brown (1984) found a different pattern of results in a recent prospective/retrospective comparison study. Subjects judged the time spent performing a word-spelling task in which they were required to rapidly recite the spelling of a series of words verbally presented to them. The subjects were assigned to one of three different versions of the spelling task which represented three levels of task difficulty. Both prospective and retrospective time judgments of the difficult condition were associated with more error than were judgments of the control and intermediate conditions. These results suggest that the expenditure of cognitive effort disrupts both prospective and retrospective timing.

To summarize, abundant evidence shows that prospective time judgments decrease as a result of increases in nontemporal processing demands. The effect of attentional demands on retrospective time judgments is less clear. Some studies suggest that intervals associated with greater amounts of cognitive processing lengthen perceived duration. Other studies, however, report opposite results or no relationship between processing demands and retrospective judgments of perceived duration (see also Bakan, 1955; Block, George, & Reed, 1980). To a large extent, a major difficulty in integrating this body of literature involves the widely differing procedures and methods used by various investigators. Different durations, tasks, time-judgment methods, and ways of analyzing the data all contribute to the apparent inconsistences that are observed. These methodological problems, in conjunction with the fact that so few paradigm-comparison studies have been reported, point to the need for more systematic, direct comparisons of timing under comparable prospective and retrospective conditions.

EXPERIMENT 1

Experiment 1 was designed to compare prospective and retrospective time judgments in a task associated with different degrees of difficutly. The experiment was also designed to address several of the difficulties noted above. First, the duration of the interval to be judged is a potentially important factor in the results. Most of the studies in this area employ intervals ranging from 10 to 120 sec, and most use only a single standard duration for all subjects to judge. In the present study, two durations from the low end of this range (16 and 32 sec) were selected because it was felt that these shorter durations might be influenced more by attentional processes and less by the memory-based processes associated with longer durations. In addition, the use of two different durations enhances the generality of the findings.

Second, the nontemporal processing tasks that subjects are required to perform differ widely across different studies. The tasks used in some studies, such as card-sorting or word-categorization, are relatively passive in nature and may represent only a weak manipulation of attentional demands. A more involving and effortful task may produce more meaningful and consistent results. The task used in the following experiment is a perceptual-motor task, chosen because it requires active participation by the subject and because the different levels of difficulty can be clearly differentiated from one another.

A final issue involves the procedures employed for analyzing time-judgment data. In most of the studies reviewed, the effects of different experimental conditions on time judgments were assessed with reference to an underestimation or an overestimation of the actual physical duration. However, an analysis based only on the direction of judgment error may fail to uncover treatment effects on subjects' responses. Michon (1972) advocated the use of variability and distribution measures on time-judgment data because such measures may be more sensitive and informative than the usual measures of central tendency. For example, one experimental condition may be associated with equal numbers of extreme underestimations and overestimations, while another condition is associated with less extreme judgments. A comparison of means or medians may lead to the erroneous conclusion that no differences exist between the two conditions and that both groups are, on the average, accurate in their judgments. Only a different type of measure, such as the absolute error of the judgments, would reveal differences that would otherwise be missed (e.g., Guay, 1982; Schwartz, 1978). Both types of error—directional and absolute—are analyzed in the present research.

Method

Subjects. Ninety-six male subjects participated in this experiment as part of a course research requirement. The subjects ranged in age from 17 to 39 years ($M = 19.2$ years). The subjects were told only that the experiment was an investigation of "perceptual processes."

Stimulus material. The stimulus consisted of a 6-pointed star figure drawn on a sheet of paper. A double boundary line formed a 4-mm-wide border around the perimeter of the figure. The figure measured 3.5 cm per side.

Procedure. The subjects were assigned randomly to the 12 treatment conditions of a $2 \times 2 \times 3$ factorial design ($n = 8$ per cell). The factors were paradigm (prospective or retrospective), duration (16 or 32 sec), and task condition (control, easy, or difficult). Each subject was tested individually.

The subjects were told that they were to perform an attentional task involving the star figure. Subjects in the control-task condition were instructed to concentrate on the figure and keep their attention focused on it until the experimenter told them to stop. Subjects in the easy-task condition were instructed to trace around the figure with a pencil, taking care to stay within the narrow border. Subjects in the difficult-task condition were given similar instructions, but were required to trace around the figure by using a mirror drawing apparatus.

Half of the subjects were tested under the prospective paradigm. These subjects were informed that they would be asked to judge the duration of the task interval after its completion and that they should monitor the time going by. The remaining subjects were tested under the retrospective paradigm and were given no information about timing.

The task interval, surreptitiously timed by the experimenter with a stopwatch, lasted either 16 sec for some subjects or 32 sec for others. The interval was initiated and terminated by the experimenter saying "start" and "stop."

Time judgments were obtained from all subjects by the methods of verbal estimation and reproduction.[1] The subjects were urged to be as accurate as possible in making these judgments. When the task interval was terminated, the subjects were first asked to provide a verbal estimate of how long a time the task appeared to last. The subjects then were asked to press a response key to illuminate a light for a duration equal to that of the task interval. The response key activated an electronic counter, located in an adjacent room, that accumulated 1-kHz pulses from a precision time base (Coulbourn Instruments S51-10) while the response key was depressed.

When the subject ended his reproduction by releasing the response key, a line printer automatically typed out the elapsed time.

Results and Discussion

The time judgments were transformed into measures representing directional error and absolute error. When different subjects judge different durations, it is a standard practice to express these measures as proportions of the duration being judged so that all scores exist on the same relative scale. For the directional measure of error, the data were converted into ratio scores by dividing each judgment by the actual elapsed time (see Hornstein & Rotter, 1969). A value less than 1 represents a judgment shorter than the actual duration, while a value greater than 1 represents a judgment longer than the actual duration. Ratio scores are typically used to assess differences in terms of a relative lengthening or shortening of perceived duration. The data were also converted into absolute error scores by subtracting the value of each judgment from the actual duration and ignoring the sign of the difference. The absolute error scores were transformed into percentages by dividing each error score by the actual duration and multiplying this value by 100; this procedure enables one to compare absolute error scores across the two durations. Absolute error scores show the proportional difference between objective clock time and judged time, and they are used to assess the overall level of accuracy of time judgments.

Each set of scores was analyzed via a split-plot factorial design which treats the two types of time judgment—verbal estimations and reproductions—as repeated measures. The two types of time judgments were significantly correlated ($r = .78$, $p < .001$). Paradigm, duration, and task condition were the between-groups factors.

Ratio score analysis. The mean time-judgment ratio scores for the various experimental conditions are presented in Table 1. An analysis of variance (ANOVA) performed on these data revealed that prospective judgments ($M = .91$) were longer than retrospective judgments ($M = .79$) [$F(1,84) = 6.81$, $p < .01$]. The duration × task interaction was also significant [$F(1,84) = 4.22$, $p < .02$]. Tests of simple main effects contrasting task conditions within each duration indicated that the source of the interaction lies in the 32-sec duration [$F(2,84) = 3.33$, $p < .05$]. Orthogonal simple comparison tests (see Keppel, 1982) were conducted to test for differences between means. These comparisons were designed to contrast the effects of no-processing and processing conditions (the control vs. combined easy-plus-difficult comparison) and to examine the effects of task difficulty (the easy vs. difficult comparison). In this analysis, the easy ($M = .92$) versus difficult ($M = .73$) comparison was significant [$F(1,84) = 5.67$, $p < .02$], indicating that increased task demands shortened perceived duration.

Table 1
Time-Judgment Ratio Scores for Experiment 1 as a Function of Group, Task Condition, and Judgment Method

	Task Condition					
	Control		Easy		Difficult	
	VE	R	VE	R	VE	R
Prospective, 16-sec Group						
Mean	.93	.92	.93	.83	.92	.84
SD	.10	.13	.18	.19	.34	.18
Prospective, 32-sec Group						
Mean	.87	.95	1.15	.98	.85	.76
SD	.29	.11	.29	.19	.38	.24
Retrospective, 16-sec Group						
Mean	.80	.78	.76	.65	1.01	.91
SD	.30	.23	.38	.23	.27	.11
Retrospective, 32-sec Group						
Mean	.87	.87	.83	.72	.61	.69
SD	.35	.22	.38	.20	.32	.20

Note—VE = verbal estimation; R = reproduction.

Table 2
Time-Judgment Percent Absolute Error Scores for Experiment 1 as a Function of Group, Task Condition, and Judgment Method

	Task Condition					
	Control		Easy		Difficult	
	VE	R	VE	R	VE	R
Prospective, 16-sec Group						
Mean	9.3	12.8	15.6	19.2	28.9	18.6
SD	6.6	6.7	10.5	16.4	17.0	15.0
Prospective, 32-sec Group						
Mean	24.6	8.9	25.7	15.5	36.7	29.5
SD	18.9	7.8	19.1	9.7	11.5	15.1
Retrospective, 16-sec Group						
Mean	27.3	23.7	39.0	36.8	18.7	10.9
SD	23.1	19.9	19.4	19.1	19.1	8.8
Retrospective, 32-sec Group						
Mean	30.8	20.2	37.1	28.9	48.8	30.3
SD	19.2	14.6	14.5	18.5	7.0	20.0

Note—VE = verbal estimation; R = reproduction.

The only other effect to achieve significance in the ANOVA was method of time judgment [$F(1,84) = 3.96$, $p < .05$], which indicated that verbal estimations ($M = .88$) tended to be longer than reproductions ($M = .83$).

Absolute error score analysis. Table 2 displays the mean absolute error scores for the various treatment conditions. An ANOVA performed on these data showed that the main effects for paradigm, duration, and task condition were significant. Less error was associated with the prospective judgments ($M = 20.5\%$) than with the retrospective judgments ($M = 29.4\%$) [$F(1,84) = 13.02$, $p < .001$]. Judgment error of the 16-sec duration ($M = 21.8\%$) was less than that of the 32-sec duration ($M = 28.1\%$) [$F(1,84) = 6.60$, $p < .01$]. The mean error scores in the main effect for task condition [$F(2,84) = 4.45$, $p < .02$] are as follows: control, 19.7%; easy, 27.3%; and difficult, 27.8%.

These effects must be interpreted with regard to the significant duration × task and paradigm × task interactions. Turning first to the duration × task interaction [$F(2,84) = 4.88$, $p < .01$], a test of simple main effects showed that there were significant differences among the task conditions within the 32-sec duration [$F(2,84) = 3.33$, $p < .05$]. Tests of orthogonal simple comparisons were applied to these data, and both comparisons were significant. There was an increase in time judgment error between the control (M = 21.1%) and combined easy-plus-difficult (M = 31.6%) conditions [$F(1,84) = 7.95$, $p < .005$], as well as between the two task conditions themselves [M = 26.8% for the easy condition, and M = 36.3% for the difficult condition; $F(1,84) = 4.92$, $p < .03$]. These results show that as the task becomes more effortful and demanding, judgments of the elapsed time tend to become increasingly inaccurate.

For the paradigm × task interaction [$F(2,84) = 4.56$, $p < .02$], an analysis of simple main effects indicated that task conditions exerted significant effects within both prospective [$F(2,84) = 5.90$, $p < .005$] and retrospective [$F(2,84) = 3.11$, $p < .05$] paradigms. Orthogonal simple comparison tests, similar to those used previously, were applied to these data. For the prospective judgments, the control (M = 13.9%) versus easy-plus-difficult (M = 23.7%) comparison was significant [$F(1,84) = 6.97$, $p < .01$], as was the easy (M = 19.0%) versus difficult (M = 28.4%) comparison [$F(1,84) = 4.82$, $p < .03$]. These results show that prospective judgments display an increasing amount of error with increases in task demands. For the retrospective judgments, neither of the orthogonal comparisons was significant. Inspection of the mean scores for the three task conditions suggests that the effect is probably due to the increase in judgment error from control (M = 25.5%) to easy (M = 35.5%) task conditions.

The main effect for time-judgment method was also significant [$F(1,84) = 12.48$, $p < .001$], indicating that more error was associated with the verbal estimations than with reproductions. However, the effect is compounded by the duration × method interaction [$F(1,84) = 4.70$, $p < .05$]. Simple main effects tests showed that the difference between verbal estimations (M = 33.9%) and reproductions (M = 22.2%) is restricted to the 32-sec duration [$F(1,84) = 16.25$, $p < .001$].

To summarize the results of Experiment 1, prospective judgments were consistently found to be relatively longer and more accurate than retrospective judgments. This finding was expected, given that prospective subjects actively attend to temporal cues, whereas retrospective subjects do not. Task conditions also had an effect on perceived duration. In general, time judgments tended to become progressively shorter and more inaccurate with increases in task demands. However, this effect was primarily confined to the 32-sec duration.

The most important feature of these results is the similarity of prospective and retrospective time judgments. Subjects in both paradigms tended to give increasingly shorter and/or inaccurate time judgments as the task interval became more difficult. The prospective results may be interpreted within the context of the attentional allocation/distraction models proposed by Thomas (Thomas & Brown, 1974; Thomas & Cantor, 1978; Thomas & Weaver, 1975), and Hicks (Hicks et al., 1977; Hicks et al., 1976; Miller et al., 1978): As nontemporal processing demands increase, less capacity is available for processing temporal cues, and a consequent decrease or unreliability in duration judgments is observed.

The data in this experiment suggest that task demands also distract retrospective subjects from processing temporal information. When unexpectedly asked to judge the duration of an interval, subjects under retrospective conditions must base their judgments on the available temporal information accumulated in memory. Assuming that a certain amount of temporal information is continually being processed and stored, even in retrospective situations, it may be that nontemporal processing demands disrupt or interfere with the usual ongoing storage of these temporal cues. Consequently, retrospective subjects in the nontemporal processing conditions must rely on temporal information of a more degraded quality for their judgments, compared with their counterparts in the control (no processing) condition.

EXPERIMENT 2

It is possible that a different aspect of attention, other than the amount of expended processing capacity, would lead to differential effects on prospective and retrospective time judgments. In particular, Underwood and Swain (1973) have argued that the deployment of attention may be an important factor. The deployment of attention refers to the extent to which attentional resources are directed to a relatively broad or narrow range of stimuli in the perceptual field (Wachtel, 1967). Underwood and Swain (1973) contend that conditions of selective (narrow) attention lengthen retrospective time judgments. In their experiment, subjects listened to 111-sec tape-recorded prose passages and attempted to detect digits embedded in the text. Since numerous studies have shown that the presence of noise enhances a selectivity or narrowing of attention, the detection task was performed under different intensities of noise background in an effort to manipulate attentional deployment. After the prose passages were presented, subjects were asked retrospectively to judge the durations of these intervals, relative to a standard interval, via the method of magnitude estimation. The passage associated with high-intensity noise was judged longer than the passage associated with low-intensity noise. In addition, fewer targets were detected in the high-noise passage, indicating that less information was processed in this condition. Thus, memory-storage-size or amount-of-processing models cannot easily account for these findings. Instead, Underwood and Swain (1973) concluded that a "greater selectivity of attention produces a longer duration experience than a broader distribution of attention" (p. 104).

The present experiment was designed to test this hypothesis in a more rigorous fashion by requiring subjects to perform selective or divided attention tasks under prospective or retrospective conditions. If Underwood and Swain's (1973) hypothesis is valid, then the selective-attention task should lengthen perceived duration, but the divided-attention task should shorten perceived duration for the retrospective subjects.

Method

Subjects. Ninety-six male subjects participated in this experiment in exchange for partial fulfillment of a course requirement. The subjects ranged in age from 18 to 34 years (M = 20.0 years). As in Experiment 1, all subjects were initially informed that the experiment was designed to study "perceptual processes."

Stimulus material. The stimuli consisted of speechlike noise presented to control subjects and dichotic word lists presented to selective- and divided-attention subjects. All stimuli were recorded on audio tape and were presented to subjects via a reel-to-reel tape recorder (TEAC Model A-2340SX) and a set of stereo headphones (Realistic NOVA 40). Both 16- and 32-sec versions of the control and dichotic tapes were constructed.

The dichotic word lists consisted of 32 word pairs for the 16-sec condition and 64 word pairs for the 32-sec condition. The word pairs were presented at a rate of 1 pair/0.5 sec. These tapes were constructed by recording a series of tones occurring at 0.5-sec intervals in channel 3. The channel 1 words were read in synchrony with the tones, and the channel 2 words were read in synchrony with those of channel 1. The result of this procedure was a dichotic tape in which the "perceptual centers" (see Fowler, 1979; Morton, Marcus, & Frankish, 1976) of each word pair were closely aligned.

The words comprising the lists were monosyllabic nouns with ratings of 5.5 or higher on the "concreteness" and "familiarity" norms of Toglia and Battig (1978). Special care was taken to ensure that the two members of each word pair were equated in terms of their familiarity ratings.

Procedure. The subjects were tested individually. As before, they were assigned randomly to one of the 12 cells of a 2×2×3 factorial arrangement. The factors were paradigm (prospective or retrospective), duration (16 or 32 sec), and task condition (control, selective attention, or divided attention). All subjects were initially presented with a short segment of speechlike noise and were asked to adjust the left- and right-channel volume controls so that the sound in each ear was of equal loudness. This procedure compensated for any differential hearing sensitivity between the two ears.

The subjects were then informed that they were to perform an attention task. Control subjects were instructed to listen carefully to a recording of speechlike noise presented on both channels. The selective- and divided-attention subjects were presented with a dichotic word list. Subjects in the selective-attention condition were instructed to listen to the word list presented in the right channel, and to ignore the list in the left channel.[2] Subjects in the divided-attention condition were instructed to attend to both right- and left-channel word lists simultaneously. Each subject was informed that his performance on the task would be assessed by a word-recognition test administered after presentation of the word list.

As in Experiment 1, subjects were assigned randomly to prospective and retrospective paradigms. The prospective subjects were told that they would be required to judge the duration of the stimulus tape at its completion; the retrospective subjects were not given this information. After presentation of the stimulus tape, verbal estimations and reproductions of the elapsed time were obtained from all subjects in the same manner as in Experiment 1.

Following the time judgments, a word-recognition test was administered to subjects in the selective- and divided-attention conditions. The recognition test was patterned after Kahneman's (1975) method for obtaining performance data on selective and divided auditory attention tasks. Subjects who had listened to the 16-sec tape (32 word pairs) were given a list of 24 words. Eight of these words had been selected randomly from the right-ear list (excluding the first and last 5 items), 8 were the corresponding left-ear words, and 8 were unpresented distractors taken from the same word pool. Subjects who had heard the 32-sec tape (64 word pairs) received a 48-word list composed of 16 right-ear words, 16 corresponding left-ear words, and 16 unpresented distractors. The words were randomized and typed on a sheet of paper. Selective-attention subjects were instructed to circle all the words they remembered hearing in the right ear; divided-attention subjects were asked to circle the words they remembered from both right and left ears.

Results and Discussion

The time-judgment data were analyzed via a split-plot factorial design, with paradigm, duration, and task condition as the between-groups factors and method of time judgment treated as a repeated measure. The correlation between verbal estimations and reproductions was .62 ($p < .001$). Time judgments were analyzed separately in terms of ratio scores and absolute error scores. In addition, selective- and divided-attention task performance was assessed via analysis of recognition test scores.

Ratio score analysis. The time-judgment ratio scores are presented in Table 3. Analysis of the ratio scores revealed, as expected, that prospective judgments (M = .87) were longer than retrospective judgments (M = .73) [$F(1,84) = 6.60$, $p < .02$]. A main effect for duration was also found, indicating that judgments of the 16-sec duration (M = .88) were proportionately longer than those of the 32-sec duration (M = .72) [$F(1,84) = 8.77$, $p < .005$].

Method of time judgment also had an effect. In contrast to Experiment 1, verbal estimations tended to be shorter than reproductions [$F(1,84) = 22.00$, $p < .001$]. However, the paradigm × method interaction was also significant [$F(1,84) = 5.18$, $p < .03$], and an analysis of simple main effects showed that verbal estimations (M = .63) were shorter than reproductions (M = .84) only within the retrospective paradigm [$F(1,84) = 24.28$, $p < .001$].

Neither the main effect for task condition ($F < 1$) nor any of the task interactions were significant, a result that

Table 3
Time-Judgment Ratio Scores for Experiment 2 as a Function of Group, Task Condition, and Judgment Method

	Task Condition					
	Control		Selective		Divided	
	VE	R	VE	R	VE	R
	Prospective, 16-sec Group					
Mean	.90	.98	.88	.88	.97	.97
SD	.19	.13	.52	.30	.37	.38
	Prospective, 32-sec Group					
Mean	.86	.96	.72	.85	.66	.77
SD	.11	.11	.34	.21	.17	.33
	Retrospective, 16-sec Group					
Mean	.54	.89	.76	.94	.84	.97
SD	.12	.33	.16	.25	.45	.51
	Retrospective, 32-sec Group					
Mean	.53	.80	.64	.78	.48	.63
SD	.23	.34	.17	.17	.10	.15

Note—VE = verbal estimation; R = reproduction.

does not support Underwood and Swain's (1973) attentional deployment hypothesis regarding the retrospective judgments. Prospective judgments were likewise unaffected by task demands.

Absolute error score analysis. The time judgments, converted into percent absolute error scores, are presented in Table 4. These data were submitted to an ANOVA which disclosed a significant main effect for paradigm [$F(1,84) = 13.70$, $p < .001$]. This effect showed that prospective judgment error (M = 26.0%) is smaller than retrospective judgment error (M = 35.4%). A significant effect for task condition [$F(2,84) = 8.66$, $p < .001$] was also found. The mean error scores for the control, selective attention, and divided attention conditions are 25.3%, 28.8%, and 37.8%, respectively. However, both main effects must be interpreted in the light of the significant paradigm × task interaction [$F(2,84) = 13.42$, $p < .001$].

Analysis of simple main effects indicated that task conditions exerted significant effects within both prospective [$F(2,84) = 13.27$, $p < .001$] and retrospective [$F(2,84) = 8.80$, $p < .001$] paradigms. For the prospective judgments, orthogonal comparison tests showed that the judgment error of the control condition (M = 25.8%) was less than that of the combined attention conditions (M = 65.1%) [$F(1,84) = 26.55$, $p < .001$]. The selective versus divided comparison was not significant. For the retrospective judgments, only the selective versus divided comparison was significant [$F(1,84) = 16.74$, $p < .001$], showing that time judgment error increased from selective attention (M = 50.4%) to divided attention (M = 86.4%) conditions.

Method of time judgment was also a significant effect in the analysis [$F(1,84) = 14.76$, $p < .001$]. As in Experiment 1, more error was associated with the verbal estimations (M = 35.2%) than with reproductions (M = 26.1%). No other effects were significant.

Table 4
Time-Judgment Percent Absolute Error Scores for Experiment 2 as a Function of Group, Task Condition, and Judgment Method

	Task Condition					
	Control		Selective		Divided	
	VE	R	VE	R	VE	R
	Prospective, 16-sec Group					
Mean	17.1	9.9	46.0	23.5	28.9	31.9
SD	12.3	8.2	21.3	21.7	21.3	18.3
	Prospective, 32-sec Group					
Mean	15.6	8.8	37.5	22.8	33.9	35.3
SD	6.6	7.0	20.8	9.2	17.7	17.0
	Retrospective, 16-sec Group					
Mean	45.3	26.2	23.4	18.3	42.1	43.2
SD	12.8	21.5	16.2	16.9	18.5	22.6
	Retrospective, 32-sec Group					
Mean	46.0	33.3	35.5	23.4	51.1	36.1
SD	23.3	18.5	17.5	14.8	10.9	15.4

Note–VE = verbal estimation; R = reproduction.

Summarizing the time-judgment data, prospective judgments were longer and more accurate than retrospective judgments. This effect is similar to that obtained in Experiment 1. The ratio scores also yielded a main effect for duration, indicating that the 16-sec interval was judged to be proportionately longer than the 32-sec interval. The absolute error scores were greatly affected by task demands. In general, time judgments tended to show increasing amounts of error as a function of task demands. Prospective judgments of the two attention conditions were virtually identical, but both were more inaccurate than those of the control condition. The retrospective judgments displayed an increasing amount of error from selective-attention to divided-attention conditions.

The retrospective data do not support Underwood and Swain's (1973) temporal-cue hypothesis insofar as the direction of error in time judgments between selective attention and divided attention conditions is concerned. However, the retrospective judgments did exhibit a significant increase in absolute error from selective- to divided-attention conditions. If the divided-attention task is viewed as being more complex and difficult than the selective-attention task, then the time-judgment data are consistent with the idea that processing demands interfere with the reception and storage of temporal information. The more that one becomes distracted from processing temporal cues, the more inaccurate one's timing becomes.

Word recognition performance. Performance on the two attention tasks was analyzed separately in terms of recognition scores devised by Kahneman (1975). Since the recognition test consisted of equal numbers of right-ear words, left-ear words, and unpresented distractors, subjects tested under selective-attention conditions are presented with half as many correct responses as are subjects tested under divided-attention conditions. Consequently, the recognition scores contain a correction factor designed to compensate for the different proportions of correct and incorrect responses on the recognition test for subjects tested under the two attention conditions.

For the selective-attention condition, the recognition score equals 2 times the number of correct responses minus the number of incorrect responses. A two-way ANOVA (paradigm × duration) performed on these scores showed no significant effects (all $Fs < 1$). For the divided-attention condition, the recognition score equals the number of correct responses minus 2 times the number of incorrect responses. These scores were submitted to a two-way ANOVA, and the main effect for paradigm was significant [$F(1,84) = 4.75$, $p < .05$]. The recognition scores of the prospective subjects (M = .81) were lower than those of the retrospective subjects (M = 3.34). This result is consistent with the notion that prospective subjects are actively attending to time-in-passing while simultaneously attempting to monitor the word lists presented to each ear.

GENERAL DISCUSSION

Two experiments assessed the role of attentional processes in prospective and retrospective time judgments. In both experiments, prospective judgments were generally longer and more accurate than retrospective judgments. This finding supports the idea that prospective subjects are motivated to actively monitor all available temporal cues, and to attend to time-in-passing. Retrospective subjects, in contrast, process temporal cues in a more incidental fashion as part of the basic information extracted from the normal ongoing flow of events. Therefore, retrospective time judgments are based on temporal information of a less precise and more degraded quality.

Retrospective subjects appeared to be processing at least some temporal information. Judgments were longer for subjects who performed the task longer; retrospective judgments also exhibited a fairly constant proportionality across the 16- and 32-sec durations. Thus, as the duration of the task interval increased, retrospective judgments of time also increased (cf. Miller et al., 1978).

Task demands also influenced time judgments, particularly in the 32-sec duration. In general, prospective judgments tended to become relatively shorter and/or more inaccurate in the presence of nontemporal task demands. These results are consistent with the outcome of a substantial number of studies and support the attentional allocation or distraction models proposed by Thomas (Thomas & Brown, 1974; Thomas & Cantor, 1978; Thomas & Weaver, 1975) and Hicks (Hicks et al., 1977; Hicks et al., 1976; Miller et al., 1978). Task demands also tended to disrupt retrospective timing. The nontemporal tasks used in these experiments were chosen because of their relevance to mental effort and attentional resources. Most previous studies were not specifically designed with this goal in mind, which may explain some of the inconsistent findings concerning nontemporal task performance and retrospective timing. The present findings suggest that nontemporal processing must be sufficiently demanding for consistent effects to be observed in retrospective judgments.

This research was also designed to try to overcome some of the methodological problems associated with previous studies in the area. One serious concern is that many studies employ only a single duration for all subjects to judge, and that different studies have used widely different durations. This fact makes it difficult to compare different studies and may explain some of the discrepancies in past research, especially when the comparison of interest is between prospective and retrospective experiments. In general, prospective studies tend to involve intervals lasting less than 60 sec, whereas most retrospective studies, focusing on memory, tend to involve intervals ranging from 60 sec to several minutes. There is some controversy over the general issue as to whether similar or different factors underlie the judgment of intervals spanning different durations. Some writers (e.g., Cohen, 1966; Ornstein, 1969, pp. 20-24) believe that judgments of different durations are mediated by different factors, and there are some data consistent with this view (Hicks & Miller, 1976; Loehlin, 1959). However, other writers (e.g., Block, 1979) believe the opposite is true, that similar factors are involved in time judgments of a wide range of durations. Thus, the issue remains unresolved. In the two experiments reported here, subjects in both paradigms judged two different durations, and duration was found to exert significant effects in almost all the analyses. This outcome suggests that duration itself may contribute to some of the conflicting results in the literature. There is a need for more parametric work in this area involving systematic comparisons of timing across different task conditions and different durations.

The issues of time-judgment method and procedures for data scoring were also evaluated. First, method of time judgment had an effect. In three of the four analyses, verbal estimations were longer and/or more inaccurate than the reproductions. These findings are consistent with those of previous research (e.g., Clausen, 1950; McConchie & Rutschmann, 1971) indicating that verbal estimations are less reliable than reproductions. It should be noted, however, that both methods produced similar findings with regard to the effects of paradigm and task conditions. This result is perhaps not too surprising, given that verbal estimations and reproductions are probably closely related (see Carlson & Feinberg, 1970; Fraisse, 1978, for discussions of this issue). Nevertheless, it is probably a good practice to include more than one judgment method in an experiment whenever possible. A related issue concerns the scoring of data and the types of information that are extracted. In this research, analysis of ratio scores (a directional measure of error) often failed to detect differences between treatment conditions, whereas the absolute error score analyses were more sensitive in detecting decrements in timing performance. The argument in favor of error and variability measures in time research (e.g., Goldstone, 1975; Michon, 1972) receives support here. Ratio scores in these experiments actually masked effects that could only be observed via analyses of absolute error. Both types of measures provide important information, and both should be included in the analysis of time-judgment data to provide a clearer picture of the effects of experimental treatments.

Perhaps the most important finding in this research concerns the basic similarity between prospective and retrospective judgments. Time judgments in both paradigms tended to be affected in the same way by manipulations of task demands. This pattern of results suggests that similar underlying processes may operate in prospective and retrospective situations. That is, nontemporal task demands serve to divert attention away from the processing and storage of temporal cues. The effect is most clearly seen in prospective conditions, since these subjects are presumably motivated to actively seek out temporal information. In retrospective conditions, nontemporal processing also appears to disrupt the accumulation of temporal cues. This interpretation is consistent with the common experience of engaging in some demanding activity for a period of time and later discovering that the objective length of the interval was much longer than one's subjective impression. In time methodology, such experiences can usually be classified as retrospective. Thus, performance of a demanding task seems to minimize temporal awareness, and this effect holds for both prospective and retrospective conditions.

REFERENCES

Allen, D. A. (1980). Filling time versus affective response to the activity that fills the time: Independent effects on time judgment? *Perceptual and Motor Skills*, **51**, 723-727.

Axel, R. (1924). Estimation of time. *Archives of Psychology*, **12**, 1-72.

Bakan, P. (1955). Effect of set and work speed on time estimation. *Perceptual & Motor Skills*, **5**, 147-148.

Block, R. A. (1979). Time and consciousness. In G. Underwood & R. Stevens (Eds.), *Aspects of consciousness* (Vol. 1), pp. 179-217. New York: Academic Press.

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

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

Brown, S. W. (1984, May). *Cognitive effort and temporal information processing*. Paper presented at the scientific meeting of the Maine Psychological Association, Bates College, Lewiston, ME.

Burnside, W. (1971). Judgment of short time intervals while performing mathematical tasks. *Perception & Psychophysics*, **9**, 404-406.

Cahoon, D., & Edmonds, E. M. (1980). The watched pot still won't boil: Expectancy as a variable in estimating the passage of time. *Bulletin of the Psychonomic Society*, **16**, 115-116.

Carlson, V. R., & Feinberg, I. (1970). Time judgment as a function of method, practice, and sex. *Journal of Experimental Psychology*, **85**, 171-180.

Clausen, J. (1950). An evaluation of experimental methods of time judgment. *Journal of Experimental Psychology*, **40**, 756-762.

Cohen, J. (1966). Subjective time. In J. T. Fraser (Ed.), *The voices of time* (pp. 257-275). New York: Braziller.

Cohen, S. (1971). Effects of task, interval and order of presentation on time estimations. *Perceptual and Motor Skills*, **33**, 101-102.

Curton, E. D., & Lordahl, D. S. (1974). Effects of attentional focus and arousal on time estimation. *Journal of Experimental Psychology*, **103**, 861-867.

DeWolfe, R. K. S., & Duncan, C. P. (1959). Time estimation as a function of level of behavior of successive tasks. *Journal of Experimental Psychology*, **58**, 153-158.

Doob, L. W. (1971). *Patterning of time*. New Haven: Yale University Press.

Fowler, C. A. (1979). "Perceptual centers" in speech production and perception. *Perception & Psychophysics*, **25**, 375-388.

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

Fraisse, P. (1978). Time and rhythm perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception. Vol. VIII: Perceptual coding* (pp. 203-254). New York: Academic Press.

Fraisse, P. (1984). Perception and estimation of time. *Annual Review of Psychology*, **35**, 1-36.

Frankenhaeuser, M. (1959). *Estimation of time*. Uppsala, Sweden: Almqvist & Wiksell.

Gilliland, A. R., Hofeld, J., & Eckstrand, G. (1946). Studies in time perception. *Psychological Bulletin*, **43**, 162-176.

Goldstone, S. (1975). The variability of temporal judgment in psychopathology. In M. L. Kietzman, S. Sutton, & J. Zubin (Eds.), *Experimental approaches to psychopathology* (pp. 393-419). New York: Academic Press.

Guay, M. (1982). Long-term retention of temporal information. *Perceptual & Motor Skills*, **54**, 843-849.

Gulliksen, H. (1927). The influence of occupation upon the perception of time. *Journal of Experimental Psychology*, **10**, 52-59.

Harton, J. J. (1938a). The influence of the difficulty of activity on the estimation of time. *Journal of Experimental Psychology*, **23**, 270-287.

Harton, J. J. (1938b). The influence of the difficulty of activity on the estimation of time: II. *Journal of Experimental Psychology*, **23**, 428-433.

Harton, J. J. (1942). Time estimation in relation to goal organization and difficulty of tasks. *Journal of General Psychology*, **27**, 63-69.

Hawkes, G. R. (1972). Time perception research: An empirical basis for the study of attention. *Psychonomic Monographs Supplement*, 4(12, Whole No. 60), 221-224.

Hawkes, G. R., & Sherman, S. J. (1972). Vigilance effects for duration judgments with two levels of task demand. *Perceptual & Motor Skills*, **34**, 351-356.

Hicks, R. E., & Brundige, R. M. (1974). Judgments of temporal duration while processing verbal and physiognomic stimuli. *Acta Psychologica*, **38**, 447-453.

Hicks, R. E., & Miller, G. W. (1976). Transfer of time judgments as a function of feedback. *American Journal of Psychology*, **89**, 303-310.

Hicks, R. E., Miller, G. W., Gaes, G., & Bierman, K. (1977). Concurrent processing demands and the experience of time-in-passing. *American Journal of Psychology*, **90**, 431-446.

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

Hornstein, A. D., & Rotter, G. S. (1969). Research methodology in temporal perception. *Journal of Experimental Psychology*, **79**, 561-564.

James, W. (1950). *The principles of psychology* (Vol. 1). New York: Dover. (Original work published 1890)

Kahneman, D. (1975). Effort, recognition and recall in auditory attention. In P. Rabbitt & S. Dornic (Eds.), *Attention and performance V* (pp. 65-80). New York: Academic Press.

Keppel, G. (1982). *Design and analysis: A researcher's handbook* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Loehlin, J. C. (1959). The influence of different activities on the apparent length of time. *Psychological Monographs*, **73**(4, Whole No. 474).

Martin, D. W. (1970). Residual processing capacity during verbal organization in memory. *Journal of Verbal Learning & Verbal Behavior*, **9**, 391-397.

McClain, L. (1983). Interval estimation: Effects of processing demands on prospective and retrospective reports. *Perception & Psychophysics*, **34**, 185-189.

McConchie, R. D., & Rutschmann, J. (1971). Human time estimation: On differences between methods. *Perceptual & Motor Skills*, **32**, 319-336.

McKay, T. D. (1977). Time estimation: Effects of attentional focus and a comparison of interval conditions. *Perceptual & Motor Skills*, **45**, 584-586.

Michon, J. A. (1972). Processing of temporal information and the cognitive theory of time experience. In J. T. Fraser, F. C. Haber, & G. W. Muller (Eds.), *The study of time* (pp. 242-258). New York: Springer.

Miller, G. W., Hicks, R. E., & Willette, M. (1978). Effects of concurrent verbal rehearsal and temporal set upon judgments of temporal duration. *Acta Psychologica*, **42**, 173-179.

Morton, J., Marcus, S., & Frankish, C. (1976). Perceptual centers (P-centers). *Psychological Review*, **83**, 405-408.

Ornstein, R. E. (1969). *On the experience of time*. Baltimore: Penguin Books.

Schwartz, W. (1978). Time and context during hypnotic involvement. *International Journal of Clinical and Experimental Hypnosis*, **26**, 307-316.

Smith, N. C. (1969). The effects on time estimation of increasing the complexity of a cognitive task. *Journal of General Psychology*, **81**, 231-235.

Sturt, M. (1925). *The psychology of time*. New York: Harcourt Brace.

Thomas, E. A. C., & Brown, I. (1974). Time perception and the filled-duration illusion. *Perception & Psychophysics*, **16**, 449-458.

Thomas, E. A. C., & Cantor, N. E. (1978). Interdependence between the processing of temporal and non-temporal information. In J. Raquin (Ed.), *Attention and performance VIII* (pp. 43-62). Hillsdale, NJ: Erlbaum.

Thomas, E. A. C., & Weaver, W. B. (1975). Cognitive processing and time perception. *Perception & Psychophysics*, **17**, 363-367.

Toglia, M., & Battig, W. (1978). *Handbook of semantic word norms*. Hillsdale, NJ: Erlbaum.

Trumbo, S., & Milone, F. (1971). Primary task performance as a function of encoding, retention, and recall in a secondary task. *Journal of Experimental Psychology*, **91**, 271-279.

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

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

Vroon, P. A. (1970). Effects of presented and processed information on duration experience. *Acta Psychologica*, **24**, 115-121.

Wachtel, P. L. (1967). Conceptions of broad and narrow attention. *Psychological Bulletin*, **68**, 417-429.

Wilsoncroft, W. E., & Stone, J. D. (1975). Information processing and estimation of short time intervals. *Perceptual & Motor Skills*, **41**, 192-194.

Zakay, D., Nitzan, D., & Glicksohn, J. (1983). The influence of task difficulty and external tempo on subjective time estimation. *Perception & Psychophysics*, **34**, 451-456.

NOTES

1. The order of obtaining the two types of time judgments was not counterbalanced. Verbal estimations were always recorded first because, in contrast to reproductions, the time required to obtain these responses is small and not dependent on the magnitudes of the response. The procedure followed here ensures a minimal and relatively constant delay between the end of the interval and the recording of the two time judgments.

2. All subjects in the selective-attention condition were instructed to attend to the word list presented to the right ear. It was not considered necessary to counterbalance right and left ears in this condition because the interstimulus interval was comparatively long (0.5 sec) and the subjects were not required to make any overt speeded responses to the stimuli as they were presented.

(Manuscript received April 22, 1985;
revision accepted for publication July 23, 1985.)