

Accuracy of temporal coding: Auditory-visual comparisons

FRANK R. SCHAB and ROBERT G. CROWDER
Yale University, New Haven, Connecticut

Three experiments were designed to decide whether temporal information is coded more accurately for intervals defined by auditory events or for those defined by visual events. In the first experiment, the irregular-list technique was used, in which a short list of items was presented, the items all separated by different interstimulus intervals. Following presentation, the subject was given three items from the list, in their correct serial order, and was asked to judge the relative interstimulus intervals. Performance was indistinguishable whether the items were presented auditorily or visually. In the second experiment, two unfilled intervals were defined by three nonverbal signals in either the auditory or the visual modality. After delays of 0, 9, or 18 sec (the latter two filled with distractor activity), the subjects were directed to make a verbal estimate of the length of one of the two intervals, which ranged from 1 to 4 sec and from 10 to 13 sec. Again, performance was not dependent on the modality of the time markers. The results of Experiment 3, which was procedurally similar to Experiment 2 but with filled rather than empty intervals, showed significant modality differences in one measure only. Within the range of intervals employed in the present study, our results provide, at best, only modest support for theories that predict more accurate temporal coding in memory for auditory, rather than visual, stimulus presentation.

The recency effect in memory captures one of the most conspicuous laws of memory and metamemory: we remember well what just happened and less well what happened in the more remote past. Currently, one prominent theoretical approach to recency is founded on the proposition that items from the end of a series have a privileged status because of their discriminability along the temporal dimension (Baddeley & Hitch, 1977; Hitch, 1985; Murdock, 1960). Glenberg (1987; Glenberg & Swanson, 1986) has set out a particularly well-formulated version of this idea. His *temporal distinctiveness* theory deals successfully with several aspects of recency including, notably, the long-term recency effect of Bjork and Whitten (1974) and its dependence on the relative spacing of list items to each other and the time lapse prior to recall.

Briefly, the central proposition of Glenberg's (1987; Glenberg & Swanson, 1986) retrieval-based distinctiveness theory is that encoding of new information includes a description of the time of occurrence of the new information. Retrieval proceeds by means of temporally defined search sets when more efficacious retrieval cues (i.e., semantic cues) are not available. The size of a given search set is a function of how long ago the information was presented: temporally more distant events are as-

sociated with a larger search set than are recent events. As the size of the search set increases, increasingly more information will tend to be consistent with the search set, and retrieval of the desired information becomes less probable. Operationally, distinctiveness thus refers to the number of encoded events consistent with a given temporal search set. A further assumption of this distinctiveness theory is that the time of occurrence of events is encoded more accurately in the auditory than in the visual modality.

This article concerns a secondary aspect of Glenberg's overall theory pertaining to this last assumption, that is, its account of the auditory/visual modality effect. The target finding is that recency following auditory presentation of linguistic items surpasses recency following visual presentation read silently (Crowder & Morton, 1969; Penney, 1975; Watkins, Watkins, & Crowder, 1974). Glenberg and Swanson (1986) proposed that this recency advantage for spoken presentation results from inherently better temporal coding for auditory than for visual input. That is, auditory items are more distinctive with regard to their times of occurrence, and therefore—since recency depends on temporal distinctiveness—show more recency than do visual items. Glenberg and Swanson (1986; also Glenberg, 1987) reported findings consistent with this principle, showing a special sensitivity of auditory presentation to timing manipulations during list presentation. However, the main dependent measure in those studies was the size of the recency effect. Although recency data represent an important component in the empirical evaluation of theoretical accounts of the recency and modality effects, such data are not entirely satisfactory without converging evidence. For example, exclusive reliance on

This research was supported by NSF Grant BNS 92-19661 to Robert G. Crowder. We wish to thank Arthur M. Glenberg, Alice F. Healy, Åke Hellström, and an anonymous reviewer for their thoughtful comments on an earlier version of this manuscript. Correspondence regarding this article should be addressed to Frank R. Schab, Department of Psychology, Yale University, P. O. Box 11-A, Yale Station, New Haven, CT 06520-7447.

recency data as evidence in support of the temporal distinctiveness theory would lead to the following, undesirable logical situation. When asked why the modality effect occurs, one would answer, "Because the auditory mode encodes time of occurrence more accurately than the visual mode does." However, when then asked about the evidence for the greater accuracy of temporal coding in audition, one would be forced to answer, "Because of the modality effect." Clearly, some independent measure of temporal distinctiveness must be used to assess whether temporal coding is, in fact, more precise in audition than in vision.

One such instance of converging evidence, which we will only mention briefly, has come from studies concerning serial-position judgments. In a series of experiments, Glenberg and Fernandez (1988, Experiments 1-3) presented subjects with lists of word pairs, in either an auditory or a visual modality, and asked the subjects to judge whether the first or second of two word pairs, which had been presented in adjacent positions in the list, had been presented more recently. Their results supported a modified version of the temporal distinctiveness theory's account of the modality effect. In a different study (Greene & Crowder, 1988, Experiments 3 and 4), subjects judged the actual serial position of list items, presented either auditorily or visually, and the results were generally consistent with the temporal distinctiveness theory.

Converging evidence of a different sort can be sought by investigating interval judgments. Temporal judgments of intervals delimited by auditory and visual stimuli are independent of recency data and provide an unbiased test of differential temporal distinctiveness in encoding between the two modalities. The argument employed here is that knowing when a series of events occurred is tantamount to knowing the extent of relative time separating the events, and vice versa. Specifically, knowledge of the times of occurrence of, say, four successively presented stimuli is akin to knowledge regarding the intervals separating them, and knowledge of the relative lengths of the intervals implies knowledge of the times of occurrence of the events. To use a spatial metaphor, if one has information concerning the relative positions of a series of points on a line, one also has information concerning the relative distances between the points, and, conversely, if one possesses the information of the relative distances between the points, one also knows where the points are relative to one another. Thus, given a perceived or remembered series of events, information about (1) when the events occurred and (2) how much time separates each event from its neighbor(s) is not independent, but complementary. Consequently, a loss of accuracy in either aspect of temporal coding—time of occurrence or interval—will result in a loss of accuracy in the other. If, for example, the times of occurrence of two events are accurately coded, then so also is the temporal interval separating those events. If, on the other hand, the two events are encoded inaccurately in terms of when they occurred, then the interval separating the events will be encoded inaccurately, too.

A possible objection to this approach may be that the temporal coding involved in time-judgment tasks, which require subjects to make temporal estimates, and the temporal coding evident in recency tasks, which require subjects to recall meaningful information, are different. This argument presupposes that the same stimuli (e.g., letters) are differentially encoded in terms of their time of occurrence, depending on whether recall or temporal judgment is going to be required. We reject this criticism for the following reasons. First, this argument is based on an unparsimonious view of information processing. Second, and more importantly, Crowder and Greene (1987, Experiment 2) tested this possibility by directing subjects to judge temporal intervals on a random half of all trials and to recall the list items on the other half. The subjects were cued for recall or time judgment only after the stimulus presentation, thereby denying them a systematic and reliable strategy. The result of their experiment was that on trials requiring recall, a modality effect was found, whereas on trials requiring temporal judgment, modality had no effect on judgment accuracy. Thus, the different-strategy argument has been inspected and rejected as incorrect.

Crowder and Greene (1987) have reported a study designed to provide an independent test of the temporal distinctiveness theory. In a series of experiments, the irregular-list technique was used to investigate temporal duration judgments. In this technique, subjects are presented with a serial list characterized by unequal temporal separation between individual list items (see Crowder, 1973). In the Crowder and Greene study, the subjects saw, on each trial, a sequence of either six letters (Experiment 1) or six common, four-letter words (Experiment 2). The six list items were separated by five different time intervals in the range of .5 to 2.5 sec. On half of the trials, the subjects were told to read the stimuli silently; on the other half, they were to read them aloud. This manipulation of presentation modality (reading aloud vs. reading silently) was chosen in view of evidence (Crowder, 1970) that reading aloud is essentially equivalent to simple auditory-only presentation. In the current article, we ask, among other things, whether this choice of manipulating input modality in the earlier study was critical for the absence of modality differences in that study.

Following stimulus presentation, the subjects in the Crowder and Greene (1987) study were given three adjacent items from the input list and were asked to judge whether the interval between the first and second members of the triplet had been greater than the interval separating the second and third, or whether the opposite was true. Thus, the irregular-list technique is an appropriate tool for detecting differences in the resolution of temporal coding between the two presentation modalities; if auditory stimuli are temporally more distinctive than visual stimuli, then the judgments of the relative intervals separating the auditory stimuli should be more accurate than the judgments of those separating the visual stimuli. In their two experiments, Crowder and Greene obtained no difference in duration judgment as a func-

tion of presentation modality. The straightforward predictions of Glenberg (1987; Glenberg & Swanson, 1986) were thus not confirmed. The present experiments were designed to explore further this relationship between temporal coding and presentation modality. However, before reporting our new studies, we pause to review the literature on temporal cognition and presentation modality.

Presentation Modality and Time Estimation

Conflicting evidence exists in the time literature concerning the effects of the modality of the stimulus defining a given temporal interval on the judged duration of that interval. Several researchers (e.g. Behar & Bevan, 1961; Goldstone, Boardman, & Lhamon, 1959; Goldstone & Goldfarb, 1964a, 1964b, 1966) have reported differential effects of auditory and visual stimuli on temporal judgments, whereas other investigators have failed to demonstrate any modality differences on duration judgments (e.g., Bobko, Thompson, & Schiffman, 1977; Hawkes, Bailey, & Warm, 1961; Hirsch, Bilger, & Deatherage, 1956; Warm, Stutz, & Vassolo, 1975) or have found inconsistent differences (e.g., Tanner, Patton, & Atkinson, 1965). In those studies that have found modality effects on time judgments (e.g., Goldstone and his associates), the investigators have generally reported an underestimation of intervals defined by visual stimuli, relative to auditory intervals of equal physical duration.

These conflicting findings are not irreconcilable, however, when their methodological differences are examined. Indeed, as most duration researchers are painfully aware, many differences in the results obtained from different time-estimation studies can be attributed to differences in methodology. In the case of the contradictory findings obtained for modality effects on duration judgments, two major methodological differences between those studies that found modality effects and those that found none must be considered: the type of time judgment required of subjects and the stimulus range employed. Typically, in those experiments in which judged duration was found to be dependent on stimulus modality, durations of less than 5 sec (mostly less than 2 sec) were employed, and their subjects were asked to make absolute judgments on the basis of subjective temporal standards, whereas in the experiments in which time judgments were found to be independent of stimulus modality, intervals ranging from 1 to 40 sec were used, and their subjects were asked to estimate or reproduce temporal intervals.

Thus, Goldstone et al. (1959), employing lights and sounds as stimuli, demonstrated that visual stimuli were overestimated relative to their subjects' internal conception of one clock second more frequently than were auditory stimuli and that, on average, auditory stimuli were judged to be longer than visual stimuli of equal physical duration. The task for the subjects in that experiment consisted of saying "more" or "less," depending on whether the subject thought the variable interval was greater or less than the standard, respectively. Subsequently, Gold-

stone and Goldfarb (1964a) presented subjects on each trial with a standard interval of 1 sec and a variable interval in the range of .6 to 1.4 sec. The intervals were marked by continuous tones or lights, and the subjects were instructed to compare the variable interval with the standard and to decide on a 5-category scale whether the variable interval was shorter, equal to, or longer than the standard. The investigators found that the auditory intervals were overestimated relative to the visual intervals. Similarly, Behar and Bevan (1961), employing a method of single stimuli with auditory and visual intervals of 1 to 5 sec, reported that when subjects judged the intervals on an 11-category scale ranging from "very, very, very short" to "very, very, very long," auditory intervals were consistently overestimated relative to visual intervals by an astounding 20%. Although Tanner et al. (1965), using a two-alternative, forced-choice paradigm with intervals ranging from .5 to 1.6 sec, found that visual intervals of .5 and .6 sec were overestimated relative to equally long auditory intervals, these researchers reported that their subjects were more accurate in making temporal comparisons when the intervals were defined by auditory rather than by visual stimuli. This finding is in accordance with that of Goodfellow (1934), who investigated subjects' sensitivity to auditory, visual, and tactile intervals of 1 sec by means of the methods of just noticeable differences, constant stimuli, and reproduction. Goodfellow's results indicated a slightly greater sensitivity to auditory intervals relative to tactile and visual intervals. Finally, in a somewhat different study, Goldstone (1968) presented sounds or lights of 1, 2, 3, and 4 sec to his subjects and asked them to reproduce those intervals. From the results of that experiment, Goldstone concluded that auditory intervals are judged longer than visual intervals of equal physical duration and that this finding reflects "a more general temporal property of the sensory systems" (p. 759). However, the range of intervals employed by Goldstone does not permit such a generalization, especially in light of the following investigations in which no modality effects on interval judgment were found.

Hirsch et al. (1956), using durations ranging from 1 to 16 sec and asking their subjects to reproduce the temporal intervals, concluded that the nature of the stimuli defining the durations—tones or lights—had no effect on interval reproductions. Hawkes et al. (1961) employed three stimulus modalities (audition, vision, and touch) and three different methods of response (production, reproduction, and verbal estimation), and found that durations in the range of .4 to 5 sec were judged with similar accuracy. Finally, Warm et al. (1975) demonstrated symmetrical duration-discrimination transfer between auditory and visual modalities with training on intervals of 6, 12, and 18 sec.

These few studies may suffice to show that those investigations supporting modality-dependence of duration judgments are methodologically different in crucial aspects from those studies demonstrating modality-independence of time judgments. For durations longer than a few sec-

onds, and for responses more precise than absolute category judgments, we suggest that both duration experience and duration judgment are mediated by central processes and are relatively independent of peripheral sensory parameters. In fact, several important theories of time discrimination (e.g., Allan & Kristofferson, 1974; Creelman, 1962; Ornstein, 1969; Treisman, 1963) have defined central processes as the operational mediators of temporal discrimination, and have maintained that duration judgments are arrived at on the basis of the temporal information inherent in the interval rather than on sensory attributes. For shorter intervals, we agree with other investigators (e.g., Fraisse, 1984) that a different judgment process may be operating, one that perhaps relies more on sensory features in which duration and temporal patterning (e.g., Garner & Gottwald, 1968) are influenced by the modality of presentation.

By way of comment on this small literature review, we note that this is one research area in which considerable work in the past has yielded almost no useful information—and this is not just because the findings are in dispute. A more serious problem is that research questions were formulated, for the most part, in terms of such variables as overestimation and underestimation. These measures, however, are inherently uninformative because they say nothing about the relative accuracy of temporal coding between the two modalities.

The goal of the present study was to investigate further the possibility that duration judgments of intervals greater than 1 sec are relatively affected by modality. Glenberg (1987; Glenberg & Swanson, 1986) has suggested that time of occurrence of stimuli is encoded more accurately when the presentation is auditory rather than visual. In accordance with the argument presented above, this auditory advantage, if correct, should hold for duration judgments as well as for time-of-occurrence judgments, since accuracy in one implies accuracy in the other. Two series of intervals were used, one composed of short intervals and the other of longer ones. The reason for using two different series of intervals is based on the assumption of the temporal distinctiveness theory that a general auditory advantage exists which is capable of explaining short-term as well as long-term recency. Furthermore, two different responses were required of subjects to provide a less restricted test of modality differences. In Experiment 1, the subjects were asked to compare two intervals, whereas in Experiments 2 and 3, the subjects were asked to make verbal estimates of one of two intervals.

EXPERIMENT 1

As noted above, Crowder and Greene (1987) failed to detect a reliable auditory-visual difference in time judgments in their studies using the irregular-list technique. A skeptic may claim that the Crowder and Greene task, requiring silent and aloud reading of printed items, does not qualify as a "pure" auditory-visual comparison. Although Crowder and Greene noted that the compari-

son of silent and vocal reading of printed lists is a sufficient condition for observing the modality effect in recall, and that the recall effect is one target of the temporal distinctiveness theory, their experiments still leave open the possibility of a true modality difference in time judgment. To counter this criticism, in Experiment 1, we pitted "pure" auditory and visual stimulus presentations against one another in a study of temporal judgment using the irregular-list technique. In an extension of Crowder and Greene, in Experiment 1, we also employed numerals as the stimuli marking the temporal intervals, rather than letters or words.

Method

Subjects. Twenty-six Yale undergraduates, volunteering in return for course credit, participated in this experiment.

Stimuli and Apparatus. The entire experiment, including stimulus presentation and response collection, was controlled by a Commodore C64 computer with an attached color monitor and a vocatrax-based speech synthesizer. The marker stimuli used in this experiment consisted of the numerals 1 through 9 and were presented either visually on the computer monitor or auditorily by the speech synthesizer, which was connected to an amplifier with a loudspeaker placed on top of the color monitor. Presentation length of the numerals was 500 msec for visual presentation and averaged 509 msec for auditory presentation. Specifically, the durations of the auditory markers were 470, 420, 500, 580, 470, 490, 610, 510, and 530 msec, respectively, for numerals 1 through 9. The five temporal intervals used in this experiment were 1.0, 1.5, 2.0, 2.5, and 3.0 sec.

Design and Procedure. Each trial consisted of the sequential presentation of the five intervals marked by six numerals chosen randomly and without replacement from the pool of nine. Each of the five intervals was paired with the other four, resulting in 20 interval combinations. One of these 20 interval combinations was tested on each trial, and the remaining three intervals of each trial were the three intervals of the original set that were not part of the test combination.

Following the presentation of the list of numerals separating the five intervals, the subjects were given the three successively presented numerals that had marked the two test intervals during trial presentation, and were asked to judge whether the interval between the first and second of those had been longer than the interval between the second and third, or whether the opposite was true. In other words, the subjects judged which of the two test intervals was longer, the first or the second. The subjects were told that the intervals began and ended with the onset of the markers. The subjects were also informed that the magnitude of the numerals marking the intervals was in no way related to the sizes of the intervals.

Each interval combination was tested four times, once in each of the four possible positions for adjacent doubles to occur within the serial list of five intervals, yielding a stimulus block of 80 trials. This block was presented both visually and auditorily, for a total of 160 trials per subject. Half of the subjects began with the visual presentation, the other half with the auditory. Trial order was always randomized within and across blocks.

Visual trials started with the words "watch here" flashed on the monitor for 1 sec. Following the removal of this preparatory message, each of the six numerals was presented for 500 msec in the center of the monitor, separated by the five temporal intervals. During the presentation of the auditory trials, the screen remained blank. Immediately after the offset of the last numeral, the subjects were prompted on the monitor with three numerals that had occurred adjacently during trial presentation and were asked to enter a "1" if they thought the time interval between the first and second of

these numerals had been longer than the interval between the second and third, and a "2" if they thought the latter interval had been longer than the former. Following the discrimination judgment, the subjects proceeded to the next trial. There was a 5-min rest interval between the presentation of the auditory and visual blocks.

Results

Figure 1 presents the discrimination results for each of the 20 interval combinations and both presentation modalities. The abscissa represents the arithmetic difference between the second and first intervals (i.e., a negative difference of -2 indicates that the first interval was 2 sec longer than the second, as in the case of the interval combination 3.0/1.0, whereas a positive difference of $+1$ indicates that the second interval was longer than the first by 1 sec, as for the following interval combinations: 1.0/2.0, 1.5/2.5, and 2.0/3.0). The ordinate represents the probability of choosing the second interval as the longer of the two intervals. The psychophysical functions approximated in Figure 1 indicate that perceived duration is a sensitive function of physical time. However, the scatterplot provides no evidence for a modality effect on duration discrimination. An analysis of variance (ANOVA) on the discrimination responses revealed no main effect for modality [$F(1,25) = 1.47$, $MS_e = .262$, $p > .05$], nor were any of the modality interactions significant. Furthermore, an examination of individual subject data showed that of the 26 subjects, 16 demonstrated greater discrimination accuracy for auditory intervals, 9 were more accurate with visual intervals, and 1 was equally accurate in both modalities ($p = .115$ by a sign test).

A different analysis of modality was performed on the subjects' pooled probability of correct discrimination for each presentation modality. The arc-sine transformed proportions were submitted to an ANOVA, which revealed no modality effect [$F(1,25) = 1.81$, $MS_e = .016$,

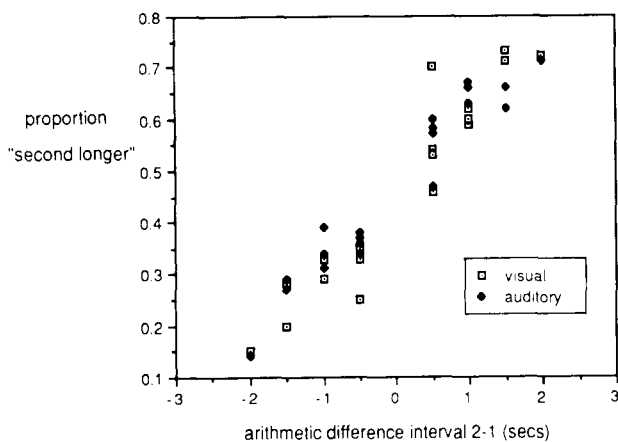


Figure 1. Probability of choosing the second interval as having been the longer of the two test intervals as a function of the arithmetic difference between Interval 2 and Interval 1 ($2 - 1$) in Experiment 1. Entries are the 20 interval combinations and the two modalities.

Table 1
Proportion of Correct Discrimination for Both Modalities and All Four Positions Within Trials in Experiment 1

Modality	Position Within Trial			
	1	2	3	4
Visual	.59	.63	.67	.78
Auditory	.57	.63	.66	.74

$p > .05$]. The mean proportions of correct discrimination were .67 for the visual mode and .65 for the auditory mode.

The mean proportion of correct discrimination judgments for auditory and visual modalities and for all four within-trial positions are shown in Table 1. If anything, the proportions indicate a slight advantage for visually marked intervals. Position was defined as the location in the serial list from which the two adjacent intervals were chosen; that is, the first and second, the second and third, the third and fourth, and the fourth and fifth intervals, respectively referred to in Table 1 as Positions 1, 2, 3, and 4. As can be seen in Table 1, discrimination accuracy, as measured by probability of correct discrimination, improved consistently from Position 1 through Position 4 [$F(3,75) = 27.17$, $MS_e = .214$, $p < .0001$]. This finding is not surprising, given the abundant evidence showing an advantage for recall of recent events over that of more distant ones. The modality \times position interaction was far from significant [$F(1,19) < 1.0$, $MS_e = .236$], however, with nearly equal proportions for auditory and visual presentations at each position.

The analysis of interval combination (each of the five intervals paired with every other interval, e.g., 1.0/1.5, 1.0/2.0, ..., 3.0/2.5) on the discrimination responses also proved to be statistically significant [$F(19,475) = 5.06$, $MS_e = .296$, $p < .01$]. A closer analysis of the 20 interval combinations was performed by examining the two factors on which the interval combinations differed: the absolute magnitude difference between the two intervals of a given combination and the direction of the difference. With regard to the absolute magnitude differences, the five intervals employed in this experiment yielded four possible magnitude differences between the two intervals tested on any given trial (.5, 1.0, 1.5, and 2.0 sec). In terms of the direction of the differences, two possibilities existed: either the first interval was longer than the second or the second was longer than the first. Table 2, which displays mean proportions of correct discrimination for absolute differences between the first and second intervals, the direction or sign of these differences, and presentation modality, shows that the subjects were more accurate when the first interval was longer than the second across all magnitude differences between the two intervals tested. In the time literature, this finding is known as a *positive time-order error*. Furthermore, there was a positive correlation between magnitude difference and discrimination accuracy: as the magnitude difference between the two intervals to be compared increased, discrimina-

Table 2
Proportion of Correct Discrimination as a Function of Modality and Directional and Magnitude Differences Between the Two Test Intervals in Experiment 1

Modality	Directional Difference	Magnitude Differences Between Intervals			
		.5	1.0	1.5	2.0
Visual	1 > 2	.68	.68	.76	.85
	2 > 1	.56	.60	.72	.72
Auditory	1 > 2	.64	.65	.72	.86
	2 > 1	.56	.65	.64	.71

tion performance improved, owing to the greater discriminability between the two test intervals. An ANOVA revealed that both the absolute magnitude difference between first and second interval [$F(3,32) = 18.3$, $MS_e = .0023$, $p < .0001$] and the direction of the magnitude difference [$F(1,32) = 24.51$, $MS_e = .0023$, $p < .0001$] were significant. Table 2 also provides no evidence for differential modality effects either on magnitude differences or on directional differences. The time-order error was true of both stimulus modalities.

A separate ANOVA was conducted to examine the possibility that the magnitude of the numerals marking the intervals may have systematically affected discrimination judgments. No such relationship was found. In terms of the relative magnitudes of the three numerals marking the two test intervals, four possible combinations existed: $a > b > c$, $a > b < c$, $a < b > c$, and $a < b < c$ (where a , b , and c represent the first, second, and third numerals, respectively). If the magnitude of the numerals had an effect on the subjects' discrimination judgments, then one would expect a decrease in discrimination accuracy on trials in which the first interval was longer than the second and the numerals marking the second interval were greater than those marking the first (i.e., "2," "4," and "9"). By the same token, a performance improvement should be expected on trials in which the numerals marking the first interval were of greater magnitude than those marking the second (i.e., "7," "3," and "1"). Of course, the reverse pattern would occur on trials in which the second interval was longer than the first. The concern about a possible effect of the numerals on duration discrimination was, however, not borne out [$F(3,2076) < 1.0$, $MS_e = .236$, for trials in which the first interval was greater than the second; $F(3,2076) < 1.0$, $MS_e = .211$, for trials in which the second interval was greater than the first].

Finally, we calculated the power of Experiment 1 to detect modality effects. We chose a conservative effect size (expected difference between the means for auditory and visual presentation) of .085 proportion correct. This difference in proportion correct was obtained in Experiment 1 for the statistically significant time-order error. Our calculation showed a power of .97 of Experiment 1 to detect a difference of .085 in proportion correct between auditory and visual modes.

Discussion

Modality. The results of this experiment provide no support for the claim that the nature of the modality—auditory or visual—of the stimuli marking intervals in the range of 1 through 3 sec has any effect on duration discrimination. It made no difference, in terms of the subjects' discrimination performance, whether the numerals defining the intervals in Experiment 1 were presented visually on the computer monitor or were spoken via the speech synthesizer. Moreover, modality failed to interact significantly with any other variable in the present experiment. Experiment 1 thus replicates Experiments 1 and 2 of Crowder and Greene (1987), which also failed to find significant effects on temporal discrimination of the modality of the stimuli delimiting the time intervals. In a direct comparison of the results of our Experiment 1 with those of Experiment 2 of Crowder and Greene, moreover, we find that the regression of the probability of choosing the second interval as the longer of the two intervals on the arithmetic difference between the second and first intervals (see Figure 1) yields similar functions. In the present experiment, we obtained regression functions of $.468 + .141x$ and $.453 + .157x$, respectively, for auditory and visual presentation, whereas Crowder and Greene, correspondingly, obtained $.454 + .130x$ and $.431 + .132x$.

A skeptic may now claim that Experiment 1 does not represent an unbiased test of modality differences in temporal coding because of the differential lengths of the auditory markers as compared with the visual ones. We believe that this is not a convincing criticism. First, temporal intervals were defined from onset to onset of the markers, and the subjects were well aware of this. Second, even in the visual modality, in which marker duration was always 500 msec, it is unrealistic to assume that each marker was processed for exactly 500 msec.

Although the modality of the markers that define an interval had no apparent effect on duration discrimination, several other factors examined in the present experiment did influence temporal discrimination performance. Thus, the absence of modality effects on time discrimination found in the present experiment may not be attributed to an insensitive design or to insufficient experimental power.

One of the factors found to affect duration discrimination was retention-interval length. As the physical duration between interval presentation and response task increased, discrimination accuracy decreased proportionately. The subjects performed worst when comparing the first and second intervals on a given trial and best when comparing the fourth and fifth intervals.

Time-order error. Experiment 1 also showed that discrimination performance is positively correlated with the difference between the to-be-discriminated intervals. This result has been obtained with a wide variety of stimuli and will not be pursued further here. More important than the greater discrimination sensitivity with larger inter-

stimulus magnitude differences is the finding of a positive time-order error in the present experiment. That is, the subjects were significantly more accurate in discriminating between the two test intervals when the first was longer than the second. A positive time-order error has been reported by many researchers (for reviews, see Doob, 1971; Fraisse, 1963; Hellström, 1985) and is a common finding for intervals of less than 4 to 5 sec. However, the nature of the time-order error, whether it is the result of response bias or perceptual sensitivity, has been a matter of long-standing dispute. Historically, the time-order error has been predominantly viewed as a perceptual-sensitivity phenomenon (e.g., Blakely, 1933; Stott, 1935; Woodrow, 1935). It was thought that there existed a critical duration, or indifference interval, toward which the internal representation of the first of two intervals gravitated. If the stimulus interval was less than the critical duration, then the first stimulus was overestimated; if it was greater, it was underestimated. More recently, however, the time-order error has been increasingly investigated in terms of response bias (e.g., Allan, Kristofferson, & Rice, 1974; Carbotte, 1973; Creelman, 1962). According to the response-bias explanation, the time-order error is dependent on the previous experiences of the individual. Simple response-bias interpretations propose that subjects prefer, for example, to respond with "first interval longer" (e.g., Engen, 1971; Luce, 1959), whereas more sophisticated bias theories attribute the time-order error to the combined effects of stimulus magnitude and an acquired decision criterion (e.g., Allan, 1977; Luce & Galanter, 1963). However, response-bias explanations of the time-order error cannot readily account for the error's dependence on factors such as interstimulus interval length and stimulus range.

Jamieson and Petrusic (1975), defending a sensitivity interpretation of the time-order error, have presented evidence that a positive error occurs when the stimulus range is restricted and a negative error occurs when the stimulus range is wide. The range of intervals used in the present experiment was fairly restricted (1 to 3 sec), and a positive time-order error was obtained. We will return to a more thorough examination of the time-order error in discussing Experiments 2 and 3, which employed a wider range of intervals.

In conclusion, although Experiment 1 replicates the results of Crowder and Greene (1987) in finding no differences in temporal discrimination between intervals marked by auditory and visual stimuli, these results do not permit us to make any generalizations regarding modality effects on duration judgments. Experiment 1 was constrained both in the range of intervals employed and in the response mode required of the subjects, and it is possible that the results of the present experiment are limited in their generality by one or both of these constraints. Experiment 2 was designed to investigate whether the absence of modality effects obtained in Experiment 1 was due to the methodological characteristics of that experiment.

EXPERIMENT 2

In Experiment 1, subjects were presented with five intervals on each trial and were asked to make a comparative-length judgment between two adjacent intervals; in Experiment 2, we abandoned the context of a memory list and employed a scaling task in which the subjects were asked to make a verbal estimate concerning the duration of a single, variable interval. In Experiment 2, we also used a wider range of intervals than in the first experiment and thus it differed from the first experiment in both response task and interval range. If Experiment 2 yielded the same results as Experiment 1, then these results could not readily be attributed to the experimental procedure. The purpose of Experiment 2 was thus to provide convergence in the test of whether the modality of the stimuli delimiting temporal intervals has an effect on the accuracy of duration judgments.

Method

Subjects. Twenty members of the Yale summer community, volunteering in return for pay, served as subjects in this experiment.

Stimuli and Apparatus. An IBM AT computer with a connected color monitor was used for stimulus presentation and response collection. The stimuli delimiting the temporal intervals consisted of a 1,000-cps tone and a cyan flash of light of 26×16 cm on the monitor, both lasting 250 msec from onset to offset. Eight intervals were used in the present experiment: 1, 2, 3, 4, 10, 11, 12, and 13 sec. Intervals 1 through 4 were called "short"; intervals 10 through 13 "long."

Design and Procedure. On each trial, the subjects were presented with two intervals, one of which was the test interval and the other an interval randomly chosen from the same category (short or long) as the test interval. The two intervals presented on any given trial were defined by three tones or flashes. Thus, the first and second tone or flash marked the first interval, and the second and third marked the second interval. Auditory and visual markers were never intermixed within trials. The computer program insured that each test interval was presented as the first and second interval of a trial equally often. Three different retention intervals (0, 9, and 18 sec) were interposed between interval presentation and response task. Both the 9- and 18-sec retention intervals were filled with a distractor task which required the subjects to solve addition problems. The addition problems were presented by the computer in the center of the screen in the form " $14 + 25 =$." Each problem remained on the screen for a total of 3 sec within which the subjects had to enter, via the keyboard, the sum of the two numbers presented. The subjects were not required to press any keys other than digit keys, and an addition problem was scored as correct only if the correct sum was entered within the allowed 3 sec. No addition problem required "carrying."

The eight intervals presented in each of two positions and followed by three retention intervals yielded a block of 48 stimulus trials, which was presented with both auditory and visual markers, for a total of 96 trials per subject. Half of the subjects began with the auditory block; the remainder began with the visual one. Trial order was randomized within each block and across subjects.

The task required the subjects to estimate how long, in seconds, the test interval was. Thus, following the retention interval on any given trial, the subjects were told to judge the duration of either the first or the second interval. Prior to the response prompt, which was presented visually on the monitor, the subjects did not know which interval was to be the estimation target. The subjects were informed that all intervals were in the range of 1 through 20 sec

and all were integer values. After the subjects entered their duration estimate on each trial, they advanced to the next trial.

Results

In addition to the verbal estimates provided by the subjects, ratio scores and absolute error scores were computed for each trial. Ratio scores were calculated by dividing verbal estimates by the true lengths of the intervals tested. The use of ratio scores was desirable because it insured that responses existed on the same relative scale. A ratio score of greater than 1 represents an overestimation, a score of less than 1 indicates an underestimation, and a score of exactly 1 means that verbal estimates and true physical length of the test interval were identical. Absolute error scores were calculated for each trial by subtracting the verbal estimates from the length of the test intervals and discarding the sign of the obtained difference. In addition, coefficients of variance were calculated for each subject, modality, and interval. The coefficient of variance is a measure of judgment reliability and was calculated by dividing the standard deviation for each subject, modality, and interval by the corresponding mean interval estimate.

Table 3 shows the means of estimations, ratio scores, absolute error scores, and coefficients of variance for all eight intervals and both modalities. Although the estimation and ratio means show some variability between the auditory and visual modalities across intervals, absolute error scores were, in general, slightly higher for visually marked intervals. This suggests that the subjects made somewhat fewer errors in their estimates of intervals delimited by tones than in those marked by flashes. However, this at best marginal effect was far from statistical significance [$F(1, 19) = 1.89, MS_e = 6.20, p > .05$] and, in light of the other nonsignificant modality results [$F < 1.0; MS_e = 12.56$ for estimation scores, $MS_e = .55$ for ratio scores, and $MS_e = .025$ for coefficients of variance], provides no reason to suspect any real difference between intervals marked by auditory and visual

stimuli. Furthermore, an examination of individual subject data revealed that of the 20 subjects, 11 showed less estimation error for the visual mode, whereas 9 were more accurate when the stimulus presentation was auditory ($p > .4$ by a sign test). Not surprisingly, however, the subjects' judgments, as measured by coefficients of variance, became less reliable as the intervals lengthened [$F(7, 133) = 12.58, MS_e = .019, p < .01$].

A different analysis¹ of modality was performed on the slopes and intercepts of the regression lines relating estimated and physical time. The slopes of these regression lines are a measure of the sensitivity of the processing mechanism to changes in physical time: as the slope increases, changes in physical time result in greater changes in judged time. If temporal coding is more accurate in the auditory modality, then a slope with greater acclivity should be obtained for auditory presentation as compared with visual presentation. For each subject and both presentation modalities, the slopes and intercepts of the regression lines were calculated and submitted to an ANOVA. The mean slopes for auditory and visual presentations were .92 and .91, respectively, and were not statistically different [$F(1, 19) < 1.0, MS_e = .004$]. The mean intercept of the regression line for auditory presentation was .56; for visual presentation, .61. The difference between the intercepts was also not significant [$F(1, 19) < 1.0, MS_e = .091$].

A significant position effect was found only for error scores [$F(1, 19) = 5.38, MS_e = .819, p < .05$], and comparison of the error means for Position 1 (1.41) and Position 2 (1.32) indicates that the subjects were more accurate when the test interval was preceded by another interval. Thus, the more recent interval was judged with greater accuracy than was the temporally more distant interval. However, no significant time-order error was obtained in Experiment 2.

Table 4 presents the means for the three retention intervals on all four response measures. The data in Table 4 show that as retention interval length increased, du-

Table 3
Estimations, Ratio Scores, Error Scores, and Variance Coefficients as a Function of Stimulus Interval and Modality in Experiment 2

	Stimulus Interval															
	1		2		3		4		10		11		12		13	
	Aud	Vis	Aud	Vis	Aud	Vis	Aud	Vis	Aud	Vis	Aud	Vis	Aud	Vis	Aud	Vis
Estimations																
<i>M</i>	1.48	1.52	2.50	2.45	3.26	3.22	4.02	4.30	10.16	9.80	10.71	10.68	11.67	11.51	12.09	12.28
<i>SD</i>	.88	.74	1.05	.95	.96	.85	.98	1.34	2.52	2.57	2.56	2.66	2.69	2.98	2.51	3.01
Ratio Scores																
<i>M</i>	1.48	1.52	1.25	1.23	1.09	1.07	1.00	1.08	1.02	.98	.97	.97	.97	.96	.93	.94
<i>SD</i>	.88	.74	.53	.48	.32	.28	.24	.33	.25	.26	.23	.24	.22	.26	.19	.23
Error Scores																
<i>M</i>	.48	.52	.58	.68	.58	.60	.68	.87	1.88	2.07	1.94	2.18	2.13	2.26	2.04	2.43
<i>SD</i>	.88	.74	1.01	.84	.81	.64	.70	1.06	1.68	1.53	1.69	1.54	1.65	2.00	1.71	1.90
Coefficients of Variance																
<i>M</i>	.33	.34	.25	.28	.20	.20	.19	.17	.12	.16	.13	.14	.12	.14	.11	.12
<i>SD</i>	.22	.22	.23	.17	.11	.11	.11	.14	.10	.12	.12	.07	.11	.15	.07	.09

Note—Aud = auditory modality, Vis = visual modality.

Table 4
 Estimations, Ratio Scores, Error Scores, and Variance Coefficients
 as a Function of Retention Interval in Experiment 2

Retention Interval	Estimations		Ratio Scores		Error Scores		Coefficients of Variance	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
0 sec	6.95	4.69	1.06	.33	1.24	1.43	.16	.14
9 sec	6.92	4.67	1.09	.46	1.42	1.62	.21	.19
18 sec	7.07	4.72	1.12	.51	1.44	1.59	.22	.19

ration estimates became longer and less accurate, and this effect was significant for all four response measures [$F(2,38) = 3.68$, $MS_e = 1.08$, $p < .05$, for estimations; $F(2,38) = 7.07$, $MS_e = .079$, $p < .01$, for ratio scores; $F(2,38) = 6.64$, $MS_e = 1.158$, $p < .01$, for absolute error scores; and $F(2,38) = 6.75$, $MS_e = .021$, $p < .004$, for coefficients of variance]. The sensitivity of the present experiment in revealing the effect of retention interval on temporal estimates thus provides some assurance that the absence of a modality difference in the present experiment was not due to an insensitive experimental design.

A contrast analysis demonstrated that the increase in estimation error as retention interval increased was mainly restricted to the presence or absence of a retention interval [$F(1,38) = 13.17$, $MS_e = 1.158$, $p < .001$]. The error increase between the 9- and the 18-sec retention intervals was not significant [$F(1,38) < 1.0$, $MS_e = 1.158$]. The subjects also performed about equally well on the distractor task on trials with 9- and 18-sec retention intervals; addition problem accuracy was 74.6% and 74.5% for the 9- and 18-sec retention periods, respectively [$t(1,278) = .08$, $p > .05$].

The short and long intervals also proved to be significantly different in terms of estimation accuracy. The subjects were considerably more accurate in estimating the short intervals in comparison to the longer ones [$F(1,19) = 49.76$, $MS_e = 21.70$, $p < .001$]. Moreover, an analysis based on ratio scores revealed that short intervals tended to be overestimated (mean ratio score = 1.21) and long intervals underestimated (mean ratio score = .97) [$F(1,19) = 63.64$, $MS_e = .452$, $p < .001$].

A separate ANOVA was performed to investigate the effects of the second, untested interval on duration estimates of the test interval. No such effects were found [$F(2,38) < 1.0$], perhaps because the untested interval was always chosen from the same length category (short or long) as the tested interval.

Discussion

Experiment 2 failed to reveal a differential effect of auditory and visual stimuli on any aspect of duration judgments, corroborating the results of Experiment 1. Accuracy of verbal estimates of intervals delimited by tones was comparable to accuracy on visually marked intervals. Moreover, modality did not interact significantly with any other factor. Apparently, verbal estimations of intervals ranging from 1 to 13 sec were independent of the modality of the stimuli marking these intervals.

Furthermore, the results of Experiment 2 demonstrate a decrease in judgment accuracy as the time between interval presentation and response prompt increased. However, in contrast to Experiment 1, in which a position effect showed that performance continuously decreased with the increased time between interval presentation and response task, a further increase in retention interval (from 9 to 18 sec) resulted in no detectable change of estimation performance. It is reasonable to assume, therefore, that the introduction of the secondary task (the addition problems) was the determining factor in producing the obtained results and, moreover, that the length of the addition task was relatively unimportant for estimation performance. The arithmetic task may have caused the decrement in estimation accuracy because addition problems and duration estimates are both based on the use of numbers. Thus, performing the addition task may have interfered with the estimation task because of their shared utilization of the numeric system.

In Experiment 2, significant differences in estimation performance for the short and long intervals were also found. The four short intervals tended to be overestimated, whereas the four long intervals were frequently underestimated. This effect is an instantiation of Hollingworth's (1910) general principle that stimulus judgments are dependent on the range of stimuli used in a given experiment such that a regression toward the mean of the stimulus range occurs on judging individual stimuli. Vierordt, in 1868, suggested such a central-tendency effect for temporal intervals, and in the time literature this effect is commonly referred to as Vierordt's law. Accordingly, Vierordt's law asserts that given a range of intervals, short intervals will be overestimated and long intervals will be underestimated as duration judgments regress toward the median value of the stimulus range. Vierordt's law has been supported by several investigations (e.g., Bobko, Schiffman, Castino, & Chiappetta, 1977; Braud & Holborn, 1966; Schiffman & Bobko, 1974; Stevens & Greenbaum, 1966) and appears to be a robust phenomenon. However, Schiffman, Bobko, and Thompson (1977), employing a magnitude-estimation task, failed to find a central-tendency effect and concluded that direct scaling procedures minimize stimulus-range effects on time judgments. In contrast, in the present experiment, using the direct scaling method of verbal estimation, we found that context affected duration estimates in a way consistent with the numerous studies supporting Vierordt's law. We suggest that a possible reason why Schiffman et al. failed

to obtain regression effects may be because these researchers examined magnitude estimates for only one physical interval (25 sec), which in their experiment was the highest, the median, and the lowest interval value in three overlapping interval series (1–25 sec, 13–37 sec, and 25–49 sec, respectively). Considering the large variability inherent in scaling tasks, as compared, for instance, with discrimination procedures, an extant central-tendency effect may have been found in the Schiffman et al. study if range effects had been investigated for each physical duration in each series of intervals.

In the context of the present paper, the importance of the obtained effects of retention interval and central tendency on duration estimates lies mainly in providing a measure of confidence that Experiment 2 was not an inherently insensitive experimental preparation. The procedures of Experiment 2 and the overall power of the experiment were sensitive enough to detect differences in duration estimates due to these variables. The finding of no modality effect on duration estimation is therefore not without meaning.

However, it is possible that in Experiment 2, using unfilled intervals allowed the subjects to adopt a counting strategy. Since the subjects knew that all intervals were integer values, they had only to keep two numbers in memory, one for each interval. Such a strategy, one might argue, would effectively prevent any modality effects from occurring. In recognition of this "strategy problem," we conducted a third experiment. The goal of Experiment 3 was to replicate the absence of modality effects on temporal estimates found in Experiment 2 employing experimental procedures that would prevent the subjects from using a counting strategy.

EXPERIMENT 3

Experiment 3 was designed to neutralize three possible criticisms of Experiment 2. First, it may be claimed that the adoption of a counting strategy by the subjects precluded the finding of modality effects. Second, since the estimation cue in Experiment 2 was always visual, Experiment 2 may not have represented an unbiased test of modality differences. Finally, the retention-interval task, like the estimation cue, was presented visually, possibly resulting in differential interference or facilitation for both modalities. Each of these three points is addressed in Experiment 3.

Method

Subjects. Twelve Yale undergraduates participated in Experiment 3 in return for course credit.

Stimuli, Design, and Procedure. Experiment 3 was identical to Experiment 2 with the following three exceptions. First, both intervals on every trial were filled with an arithmetic task. With the onset of each new interval, a 4×4 matrix of two-digit numbers in the range of 10 through 40 appeared in the center of the monitor. The subjects were asked to add adjacent numbers together and speak the result out loud. The subjects' responses on the arithmetic task were taped, and they were informed of this. We did not,

however, examine accuracy on the arithmetic task. The numbers to be added within any given matrix contained "easy" pairs (i.e., "10" and "25") as well as "difficult" pairs (i.e., "39" and "17"). Therefore, it is unlikely that the subjects could adopt a strategy of rhythmic addition and using the number of addition problems solved as an estimate of the interval. Second, the two retention intervals of Experiment 1 were dropped, and the subjects made their estimates immediately following the second interval. Finally, the cue informing the subjects which of the two intervals to estimate on each trial was either auditory or visual in nature. The auditory cue was a single tone of 1-sec duration. If the tone was high-pitched (1600 Hz), the subjects were directed to estimate the first interval of the pair; if the tone was low-pitched (200 Hz), the subjects estimated the second interval. The visual cue was a small (3×3 cm), colored square presented on the computer monitor for 1 sec. If the square was presented at the top of the screen, the subjects estimated the first interval; if it appeared on the bottom of the screen, they estimated the second interval. Whether the cue informing the subjects which interval to estimate was auditory or visual was randomly determined by the computer program, with the constraint that an equal number of auditory and visual cues be used on trials in which the intervals were marked by either modality.

Results and Discussion

Modality. Similar to Experiment 2, four measures of the subjects' performance were considered: estimates, absolute error scores, ratio scores, and coefficients of variance. Table 5 presents the mean estimations, ratio scores, absolute errors, and coefficients of variance as a function of interval length and modality. With the exception of the 1-sec interval, auditorily marked intervals were estimated as having been longer than intervals delimited by visual stimuli: the mean interval estimations were 6.00 sec for auditory trials and 5.53 sec for visual trials. This difference nearly reached statistical reliability [$F(1,11) = 3.96$, $MS_e = 10.65$, $p < .073$] and is a typical finding in the time literature (e.g., see Goldstone and his associates). Interestingly, whether the overestimation of auditory intervals relative to visual intervals (or the relative underestimation of visual intervals) was the result of differential sensitivity or bias toward intervals bounded by auditory and visual stimuli is still a largely unexplored question. Also, the mean estimations of Experiment 3 were lower than those of Experiment 2, owing largely to the interpolated task. Previous research (e.g., Hicks, Miller, & Kinsbourne, 1976) has shown that as the processing demands during an interval increase, the estimates of the interval become shorter. More important to our present intentions, however, is that the various measures of judgment accuracy, and the means for ratio scores, absolute errors, and coefficients of variance in Table 5, show no consistent differences between auditory and visual modalities. The mean estimation error for intervals marked by tones was 2.35 sec and for visually marked intervals, 2.42 sec [$F(1,11) < 1.0$, $MS_e = 8.36$]. Ratio scores, too, demonstrated no modality differences [$F(1,11) < 1.0$, $MS_e = 1.01$, with means of 1.08 and 1.05 for auditory and visual trials, respectively]. Similarly, modality of interval markers had no effect on coefficients of variance. The mean coefficient of variance

Table 5
Estimations, Ratio Scores, Error Scores, and Variance Coefficients as a
Function of Stimulus Interval and Modality in Experiment 3

	Stimulus Interval															
	1		2		3		4		10		11		12		13	
	Aud	Vis	Aud	Vis	Aud	Vis	Aud	Vis	Aud	Vis	Aud	Vis	Aud	Vis	Aud	Vis
Estimations																
<i>M</i>	1.81	2.19	2.83	2.60	3.56	3.38	4.35	4.02	8.25	7.73	8.98	7.85	8.69	7.85	9.50	8.60
<i>SD</i>	1.51	2.09	1.40	1.09	1.98	1.31	1.73	1.68	2.71	2.32	3.47	2.74	3.10	3.14	3.78	2.99
Ratio Scores																
<i>M</i>	1.82	2.19	1.42	1.30	1.19	1.12	1.09	1.01	.83	.77	.82	.71	.72	.65	.73	.66
<i>SD</i>	1.51	2.09	.70	.54	.66	.44	.43	.42	.27	.23	.32	.25	.26	.26	.29	.23
Error Scores																
<i>M</i>	.81	1.19	1.04	.73	1.23	.96	1.35	1.23	2.71	2.65	3.40	3.56	4.33	4.44	4.33	4.65
<i>SD</i>	1.51	2.09	1.25	1.01	1.64	.97	1.12	1.13	1.73	1.87	2.11	2.15	2.75	2.70	2.75	2.57
Coefficients of Variance																
<i>M</i>	.30	.46	.25	.29	.26	.24	.20	.25	.21	.22	.22	.25	.25	.32	.27	.20
<i>SD</i>	.23	.33	.12	.12	.07	.15	.12	.12	.15	.12	.12	.10	.10	.15	.14	.09

Note—Aud = auditory modality, Vis = visual modality.

for auditory trials was .25 and for visual trials, .28 [$F(1,11) = 2.24$, $MS_e = .022$, $p > .16$]. Finally, none of the modality interactions approached statistical significance.

An analysis was also performed on the slopes and intercepts of the regression lines relating physical and estimated time. As in Experiment 2, regression slopes and intercepts were calculated for each subject and both modalities. The mean interval estimates for both modalities with the regression functions are shown in Figure 2. As Figure 2 shows, the slope for auditory presentation (.632) was closer to unity than was the slope for visual presentation (.545), and this difference was statistically significant [$F(1,11) = 8.02$, $MS_e = .006$, $p < .016$]. No modality difference was found in the magnitude of the regression intercept: 1.57 and 1.71, respectively, for auditory and visual modes [$F(1,11) < 1.0$, $MS_e = .516$].

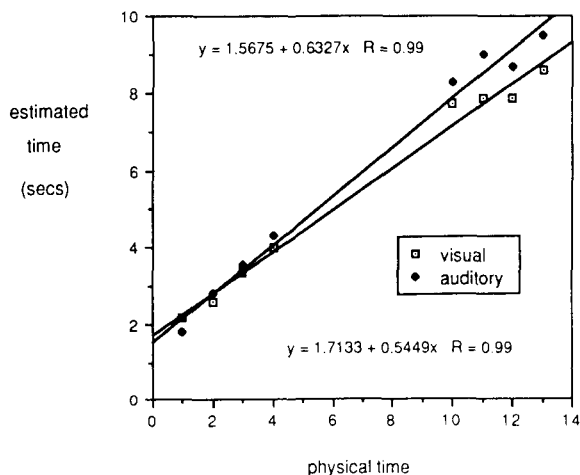


Figure 2. Mean interval estimates as a function of interval length and presentation modality in Experiment 3. The lines represent the least-squares regression functions for estimates in both modalities.

Thus, in contrast to Experiment 2, the results of Experiment 3 revealed that the subjects were more sensitive to differences in physical time when interval markers were auditory rather than visual. This finding is consistent with the temporal distinctiveness theory and will be discussed more fully below.

A possible concern in Experiment 2 was that the modality of the cue informing the subjects which interval to estimate may have adversely affected the subjects' performance. In Experiment 3, we tested this possibility by using both auditory and visual estimation cues. This concern, however, was not borne out [$F(1,11) = 1.51$, $MS_e = .58$, $p > .24$, for error scores; $F(1,11) = 1.42$, $MS_e = .026$, $p > .25$, for variance coefficients; and $F(1,11) < 1.0$ for both estimations and ratio scores]. The interaction of modality of interval markers and modality of estimation cue did not approach statistical significance for any performance measure ($F < 1.0$ for all measures).

Time-order error. Filling the intervals with the arithmetic task additionally produced a large time-order error. Intervals presented as the first of the pair were underestimated (mean = 5.32 sec) relative to intervals presented second (mean = 6.20 sec) [$F(1,11) = 12.52$, $MS_e = 11.88$, $p < .005$]. Thus, in contrast to Experiment 1, in Experiment 3 a negative time-order error was obtained. Intervals presented first were also associated with greater error (mean = 2.78 sec) than were intervals presented second (mean = 2.00 sec) [$F(1,11) = 13.98$, $p < .0034$]. Figure 3 shows mean estimations as a function of physical interval length and presentation position, and the lines represent the best-fitting linear regression functions for both presentation positions. Schab and Crowder (1988) have reasoned that analysis of the slopes and intercepts of the regression lines fit to the interval estimates for each presentation position provides information about the source of the time-order error: the slope of such a regression function measures how "fast" estimated time

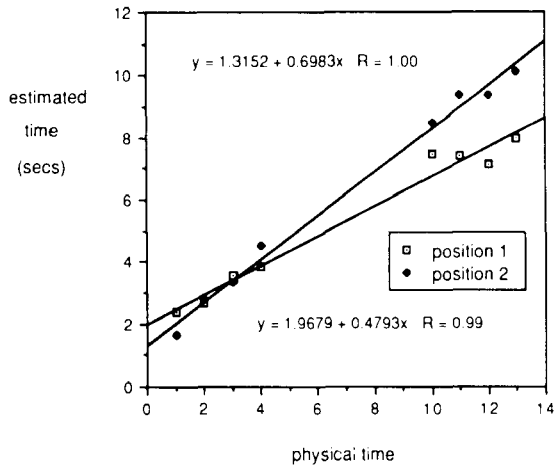


Figure 3. Mean interval estimates as a function of interval length and presentation position in Experiment 3. The lines represent the least-squares regression functions for estimates in the two presentation positions.

increases as a function of increases in physical time, and is thus a measure of sensitivity. The intercept of the regression line, on the other hand, represents bias in this context. If the regression lines fit to the estimations for each position differ primarily in intercept and have essentially parallel slopes, then the subjects are equally sensitive to increases in physical time regardless of presentation position, and the time-order error is the result of a bias on the part of the subjects to over- or underestimate an interval presented first, relative to one of the same duration presented second. If, however, the regression functions differ principally in slope (with or without a difference in intercept), then the time-order error is due to differential sensitivity of the processing system to intervals presented sequentially. In the present case, we calculated the regression slopes and intercepts for each subject and presentation position and submitted them to an ANOVA. Mean slopes differed significantly as a function of position [.479 vs. .698, respectively, for Positions 1 and 2; $F(1,11) = 14.41$, $MS_e = .02$, $p < .003$]. Intercept values for each position also differed (1.97 for Position 1 and 1.32 for Position 2), though the difference was not as pronounced as that for slopes and was only marginally significant [$F(1,11) = 5.74$, $MS_e = .44$, $p < .04$]. In conclusion, the time-order error obtained in Experiment 3 was due to the differential sensitivity to serial position of the cognitive machinery engaged when estimating the duration of intervals.

Unfortunately, the positive time-order error obtained in Experiment 1 cannot readily be submitted to the regression analysis employed here, and we can thus only speculate about its source. Experiment 3, employing a relatively wide range of intervals (1 to 13 sec), yielded a predominantly negative time-order error, whereas Experiment 1, using a fairly restricted range of intervals (1 to 3 sec), showed a positive error. This result is in accordance with that of Jamieson and Petrusic (1975). It is pos-

sible that bias enters temporal judgment more severely when the range of intervals is narrow, but whether such a hypothesis is true remains for further and direct testing. The methodological differences between Experiments 1 and 3, however, preclude an investigation of this point in the present study.

GENERAL DISCUSSION

The three experiments described here were designed to investigate whether the stimulus modality of the markers defining temporal intervals greater than 1 sec has an effect on duration judgments. As discussed above, Glenberg (1987; Glenberg & Swanson, 1986) suggested that auditory events are, in general, more accurately, or more distinctly, encoded in terms of their time of occurrence than are visual events. Unfortunately, the bulk of evidence adduced for the temporal distinctiveness theory is recall data which, in our opinion, is insufficient as evidence for a theory of recall. Converging evidence is necessary for the evaluation of the distinctiveness theory. We have argued in this paper that distinct encoding of the relative time of occurrence of a series of stimuli implies distinct encoding of the interstimulus intervals, and vice versa. Our logic was that if the times of occurrence of auditory stimuli are more distinctly encoded than those of visual stimuli, as Glenberg suggested, then so will the intervals between auditory stimuli be more accurately encoded. By inference, if subjects are more accurate in judging time intervals marked by auditory, rather than by visual, events, we can conclude that the times of occurrence of auditory events are more precisely encoded than are those of visual ones. Such a finding would provide evidence for Glenberg's proposition of greater temporal distinctiveness in the encoding of auditory, as compared with visual, events that is independent of recall data.

The results of the present experiments provide only limited support for such modality effects. In Experiment 1, the subjects discriminated between two intervals in the range of 1 to 3 sec, marked by visually or auditorily presented numerals. Discrimination accuracy was not affected by the modality of the interval markers. In Experiments 2 and 3, subjects were required verbally to estimate the length of intervals defined by tones and lights, ranging from 1 to 4 sec and from 10 to 13 sec. Estimation performance in Experiment 2, like discrimination accuracy in Experiment 1, was independent of the modality of stimuli delimiting physical time. Experiment 3 generally replicated the results of Experiment 2 but showed a significant modality difference by one performance measure—the slopes of the regression lines relating physical and estimated time. The steeper slope obtained for auditory over visual presentation implies that the subjects were more sensitive to changes in physical time when the intervals were marked by tones rather than by lights. This greater sensitivity to temporal manipulations in the auditory mode is consistent with the distinctiveness theory. At the same time, however, Experiments

1 and 2 failed to reveal any modality effects, by any measure, on temporal judgments, and even in Experiment 3, four other measures failed to show modality differences.

As in so much empirical literature with partially divergent and partially convergent results, what one chooses to emphasize depends on one's prior beliefs. Between the experiments reported here and those reported by Crowder and Greene (1987), five investigations have been reported on how well people remember the timing of auditory events.² Four of these five experiments provide no comfort whatsoever to the idea that we remember the timing of auditory events better than that of visual events. The fifth experiment (the present Experiment 3) does support that hypothesis, if one chooses only one out of the five response measures used. A true believer would understandably seize on this last finding as being the critical one, the more so because technical difficulties can be found in the design of some of the earlier studies and because the slope measure is perhaps the best motivated index of sensitivity to time. The more conservative conclusion is that the case for Glenberg's hypothesis is not yet solidly documented although, at the same time, it is not entirely without support. Perhaps our main claim at this point is that tests of this hypothesis absolutely demand experiments that measure time sensitivity directly, rather than inferring it from recall or from memory for order. We believe, in conclusion, that the burden of evidence still rests squarely on those who wish to advance this hypothesis.

As a final remark, we do not wish to imply here that auditory-visual differences in temporal coding do not exist, but only that they are small, at most, for intervals that are crucial to obtaining recency effects in memory.

REFERENCES

- ALLAN, L. G. (1977). The time-order error in judgments of duration. *Canadian Journal of Psychology*, **31**, 24-31.
- ALLAN, L. G., & KRISTOFFERSON, A. B. (1974). Psychophysical theories of duration discrimination. *Perception & Psychophysics*, **16**, 26-34.
- ALLAN, L. G., KRISTOFFERSON, A. B., & RICE, M. E. (1974). Some aspects of perceptual coding of duration in visual duration discrimination. *Perception & Psychophysics*, **15**, 83-88.
- BADDELEY, A. D., & HITCH, G. J. (1977). Recency revisited. In S. Dornic (Ed.), *Attention and Performance* (Vol. 6, pp. 647-667). Hillsdale, NJ: Erlbaum.
- BEHAR, I., & BEVAN, W. (1961). The perception of duration of auditory and visual intervals: Cross-modality comparisons and interactions. *American Journal of Psychology*, **74**, 17-26.
- BJORK, R. A., & WHITTEN, W. B. (1974). Recency-sensitive retrieval processes in long-term free recall. *Cognitive Psychology*, **6**, 173-189.
- BLAKELY, W. (1933). *The discrimination of short empty temporal intervals*. Unpublished doctoral dissertation, University of Illinois, Champaign-Urbana, IL.
- BOBKO, D. J., SCHIFFMAN, H. R., CASTINO, R. J., & CHIAPPETTA, W. (1977). Contextual effects in duration experience. *American Journal of Psychology*, **90**, 577-586.
- BOBKO, D. J., THOMPSON, J. G., & SCHIFFMAN, H. R. (1977). The perception of brief temporal intervals: Power functions for auditory and visual stimulus intervals. *Perception*, **6**, 703-709.
- BRAUD, W. G., & HOLBORN, S. W. (1966). Temporal context effects with two judgmental languages. *Psychonomic Science*, **6**, 151-152.
- CARBOTTE, R. M. (1973). Retention of time information in forced-choice duration discrimination. *Perception & Psychophysics*, **14**, 440-444.
- CREELMAN, C. D. (1962). Human discrimination of auditory duration. *Journal of the Acoustical Society of America*, **34**, 582-593.
- CROWDER, R. G. (1970). The role of one's own voice in immediate memory. *Cognitive Psychology*, **1**, 157-178.
- CROWDER, R. G. (1973). The delayed stimulus suffix effect following arhythmic stimulus presentation. *Quarterly Journal of Experimental Psychology*, **25**, 433-439.
- CROWDER, R. G., & GREENE, R. L. (1987). On the remembrance of times past: The irregular list technique. *Journal of Experimental Psychology: General*, **116**, 265-278.
- CROWDER, R. G., & MORTON, J. (1969). Precategorical acoustic storage (PAS). *Perception & Psychophysics*, **5**, 365-373.
- DOOB, L. W. (1971). *The patterning of time*. New Haven, CT: Yale University Press.
- ENGEN, T. (1971). Psychophysics: 1. Discrimination and detection. In L. A. Kling and J. W. Riggs (Eds.), *Woodworth & Schlosberg's experimental psychology* (pp. 11-46). New York: Holt, Rinehart, & Winston.
- FRAISSE, P. (1963). *The psychology of time*. New York: Harper & Row.
- FRAISSE, P. (1984). Perception and estimation of time. *Annual Review of Psychology*, **35**, 1-36.
- GARNER, W. R., & GOTTFELD, R. L. (1968). The processing and learning of temporal patterns. *Quarterly Journal of Experimental Psychology*, **20**, 97-109.
- GLENBERG, A. M. (1987). Temporal context and memory. In D. S. Gorfein & R. R. Hoffman (Eds.), *Memory and learning: The Ebbinghaus centennial conference* (pp. 173-190). Hillsdale, NJ: Erlbaum.
- GLENBERG, A. M., & FERNANDEZ, A. (1988). Evidence for auditory temporal distinctiveness: Modality effects in order and frequency judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 728-739.
- GLENBERG, A. M., & SWANSON, N. G. (1986). A temporal distinctiveness theory of recency and modality effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **12**, 3-15.
- GOLDSTONE, S. (1968). Production and reproduction of durations: Intersensory comparisons. *Perceptual & Motor Skills*, **26**, 755-760.
- GOLDSTONE, S., BOARDMAN, W. K., & LHAMON, W. T. (1959). Intersensory comparisons of temporal judgments. *Journal of Experimental Psychology*, **57**, 243-248.
- GOLDSTONE, S., & GOLDFARB, J. L. (1964a). Auditory and visual time judgment. *Journal of General Psychology*, **70**, 369-387.
- GOLDSTONE, S., & GOLDFARB, J. L. (1964b). Direct comparison of auditory and visual durations. *Journal of Experimental Psychology*, **67**, 483-485.
- GOLDSTONE, S., & GOLDFARB, J. L. (1966). The perception of time by children. In A. H. Kidd & J. L. Rivoire (Eds.), *Perceptual development in children* (pp. 445-486). New York: International Universities Press.
- GOODFELLOW, L. D. (1934). An empirical comparison of audition, vision, and touch in the discrimination of short intervals of time. *American Journal of Psychology*, **46**, 243-258.
- GREENE, R. L., & CROWDER, R. G. (1988). Memory for serial position: Effects of spacing, vocalization, and stimulus suffixes. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **14**, 740-748.
- HAWKES, G. R., BAILEY, R. W., & WARM, S. S. (1961). Method and modality in judgments of brief stimulus duration. *Journal of Auditory Research*, **39**, 345-388.
- HELLSTRÖM, Å. (1985). The time-order error and its relatives: Mirrors of cognitive processes in comparing. *Psychological Bulletin*, **97**, 35-61.
- 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.
- HIRSCH, I. J., BILGER, R. C., & DEATHERAGE, B. H. (1956). The effect of auditory and visual background on apparent duration. *American Journal of Psychology*, **69**, 561-574.
- HITCH, G. J. (1985). Short-term memory and information processing in humans and animals: Towards an integrative framework. In L. G. Nilsson & T. Archer (Eds.), *Perspectives on learning and memory* (pp. 119-136). Hillsdale, NJ: Erlbaum.

- HOLLINGWORTH, H. L. (1910). The central tendency of judgment. *Journal of Philosophical Psychology*, **7**, 461-469.
- JAMIESON, J. D., & PETRUSIC, W. M. (1975). The dependence of time-order error direction on stimulus range. *Canadian Journal of Psychology*, **29**, 175-182.
- LUCE, R. D. (1959). *Individual choice behavior*. New York: Wiley.
- LUCE, R. D., & GALANTER, E. (1963). Discrimination. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 191-243). New York: Wiley.
- MURDOCK, B. B., JR. (1960). The distinctiveness of stimuli. *Psychological Review*, **67**, 16-31.
- ORNSTEIN, R. E. (1969). *On the experience of time*. Baltimore: Penguin.
- PENNEY, C. G. (1975). Modality effects in short-term verbal memory. *Psychological Bulletin*, **82**, 68-84.
- SCHAB, F. R., & CROWDER, R. G. (1988). The role of succession in temporal cognition: Is the time-order error a recency effect of memory? *Perception & Psychophysics*, **44**, 233-242.
- SCHIFFMAN, H. R., & BOBKO, D. J. (1974). Effects of stimulus complexity on the perception of brief temporal intervals. *Journal of Experimental Psychology*, **103**, 156-159.
- SCHIFFMAN, H. R., BOBKO, D. J., & THOMPSON, J. G. (1977). The role of stimulus context on apparent duration. *Bulletin of the Psychonomic Society*, **10**, 484-486.
- STEVENS, S. S., & GREENBAUM, H. B. (1966). Regression effect in psychophysical judgment. *Perception & Psychophysics*, **1**, 439-446.
- STOTT, L. H. (1935). Time-order errors in the discrimination of short tonal durations. *Journal of Experimental Psychology*, **18**, 741-766.
- TANNER, T. A., JR., PATTON, R. M., & ATKINSON, R. C. (1965). Intermodality judgments of signal duration. *Psychonomic Science*, **2**, 271-272.
- TREISMAN, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock." *Psychological Monographs*, **77**, 1-31.
- VIERORDT, K. (1868). *Der Zeitsinn nach Versuchen*. Tübingen, West Germany: H. Laupp.
- WARM, J. S., STUTZ, R. M., & VASSOLO, P. A. (1975). Intermodal transfer in temporal discrimination. *Perception & Psychophysics*, **18**, 281-286.
- WATKINS, M. J., WATKINS, O. C., & CROWDER, R. G. (1974). The modality effect in free and serial recall as a function of phonological similarity. *Journal of Verbal Learning & Verbal Behavior*, **13**, 430-447.
- WOODROW, H. (1935). The effect of practice upon time-order errors in the comparison of temporal intervals. *Psychological Review*, **42**, 127-152.

NOTES

1. We wish to thank Åke Hellström for suggesting this analysis.
2. The experiments of Glenberg and Fernandez (1988) and Greene and Crowder (1988) concern the relative order of events, not necessarily their timing.

(Manuscript received June 29, 1987;
revision accepted for publication December 11, 1988.)