

Aging and judgments of duration: Effects of task complexity and method of estimation

FERGUS I. M. CRAIK and JANINE F. HAY

*University of Toronto, Toronto, Ontario, Canada
and Rotman Research Institute, Toronto, Ontario, Canada*

The effects of aging on judgments of short temporal durations were explored using the prospective paradigm and the methods of verbal estimation and production. Younger and older adults performed a perceptual judgment task at five levels of complexity for periods of 30, 60, and 120 sec. Participants either continued to perform the task for a specified interval (production) or were stopped and then verbally estimated the interval. Older adults gave shorter verbal estimates and longer productions than did younger adults. The methods of verbal estimation and production yielded approximately equal duration judgment ratios once range effects were taken into account. Task complexity had little effect. The major conclusion is that duration judgment ratios decrease from younger to older adults when the intervals are filled with a mental task.

Older adults generally report that time appears to pass at a faster rate at their present age than it did when they were younger (Fraisse, 1984; Schroots & Birren, 1990). Some evidence supporting this point was presented by Lemlich (1975), who suggested that the subjective duration of a year varies inversely with the square root of the person's age and reported results in line with his hypothesis. Lemlich's results were confirmed in general by Walker (1977) and by Joubert (1983). It might therefore be concluded that some biological clock (or some complex of interrelated clocks; Schroots & Birren, 1990) slows progressively as a function of increasing age, thereby giving rise to the subjective feeling that external time is passing ever faster. However, present-day researchers in the area of time perception are in good agreement that theories of human duration judgment couched solely in terms of clocks or pacemakers are inadequate (Block, 1990; Friedman, 1990; Michon, 1990; Zakay & Block, 1996). The main reason for this consensus is that judgments of duration are affected by a number of cognitive variables—such as the estimation method, environmental conditions, information processing load during the interval to be judged, the importance of the judgment to the participant, and the participants' expectations (Zakay & Block, 1996)—and it is difficult to see how such variables would affect a purely biological mechanism. Alternatively, the experience of time passing has been viewed

as a cognitive construction, as “a derived and highly formal product of the mind” (Michon, 1990, p. 38). Thus, judgments of duration depend in part on the amount of attention paid to the passage of time and temporal information processing (e.g., Hicks, Miller, & Kinsbourne, 1976; Zakay & Block, 1996), the stored contents of memory for the judged interval (Block & Zakay, 1997; Ornstein, 1969), the presence of temporal markers segmenting the judged interval (Poynter, 1983), and the number of external and internal contextual changes occurring within a time interval (Block & Reed, 1978).

Some current models embrace both biological and psychological factors, and this may constitute a reasonable middle ground. Examples of such mixed models include the ideas on rhythmic attending proposed by Jones and Boltz (1989), in which the perceiver attunes his or her rhythmic expectancies to the dynamic structure of external events, and the attentional-gate model (Block & Zakay, 1996; Zakay & Block, 1996), which augments the idea of an internal clock or counter (Treisman, 1963) with a role for the amount of attention paid to time passing. In greater detail, the attentional-gate model postulates a cognitive mechanism (the gate) that opens more widely or more frequently as more attention is paid to time; in turn, this increased influx of temporal information causes a pacemaker to transfer more pulses to a cognitive counter. Duration judgments are then based on the number of pulses that have accumulated in the counter (Zakay & Block, 1996). A mixed model of this type is attractive in that it allows for the effects of both biological and cognitive variables on the final judgment of duration.

On the biological side, there is reasonably consistent evidence that CNS depressants are associated with reductions in verbal estimations of time intervals and thus arguably with the slowing of the internal clock, whereas CNS stimulants have the opposite effect (Friedman, 1990; Meck, 1996). For example, depressants such as secobar-

The experiments reported here were supported by a grant from the Natural Sciences and Engineering Research Council of Canada to F.I.M.C. A preliminary report of this research was presented at the Cognitive Aging Conference in April 1994. The authors wish to thank Lorraine Allan and Richard Block for very helpful comments on an earlier version of this manuscript. The authors also thank Randy Sollenberger and Lianne Carley for technical help. Correspondence should be addressed to F. I. M. Craik, Department of Psychology, University of Toronto, Toronto, ON M5S 3G3, Canada (e-mail: craik@psych.utoronto.ca).

bital (Goldstone & Kirkham, 1968), ethanol (Tinklenberg, Roth, & Koppell, 1976), and cyclopropane (Adam, Rosner, Hosick, & Clark, 1971) are associated with reductions in verbal estimates of duration, whereas stimulants such as amphetamines (Church, 1984; Goldstone & Kirkham, 1968) and marijuana (Hicks, Gualtieri, Mayo, & Perez-Reyes, 1984; Tinklenberg et al., 1976) are associated with increased verbal estimates of duration, implying that the internal clock is running too fast. On the cognitive side, there is good evidence that more demanding or more complex tasks are associated with shorter verbal estimates of duration in the prospective paradigm, in which subjects know in advance that they will be asked to estimate the time interval (Block & Zakay, 1997; Hicks et al., 1976).

What do these various theories predict with respect to aging? The observation by older adults that time now appears to pass at a faster rate suggests the slowing of some internal clock or pacemaker, although that notion may be too simplistic (see Schroots & Birren, 1990, for a thoughtful discussion). It also seems likely that judgments about the rate at which months and years appear to pass are made on a different basis from judgments about intervals in the seconds-to-minutes range. Nevertheless, the evidence from psychopharmacological studies suggests that CNS depressants are associated with verbal underestimations of duration (smaller ratios of subjective to objective duration), whereas CNS stimulants are associated with verbal overestimations (larger duration judgment ratios), and these results are in line with the reports of older people on the assumption that CNS depressants mimic the effects of aging in some respects at least. Thus, the reasonable prediction from a biological clock perspective is that the clock or pacemaker slows with advancing age, resulting in smaller duration judgment ratios in the elderly than in their younger counterparts.

From a cognitive perspective, a decrease in duration judgment ratios is also the most reasonable expectation. If judgments of duration depend heavily on the amount of attention paid to the passage of time, older people arguably have fewer attentional resources to deploy (e.g., Craik & Byrd, 1982), resulting in less frequent openings of the attentional gate, fewer accumulated pulses in the cognitive counter, and, therefore, shorter verbal estimates of the temporal interval. This predicted association of aging and shorter duration judgment ratios should be stronger in dual-task situations, given the evidence that older participants are more negatively affected by the necessity to perform two tasks at once (Craik, 1977; McDowd & Craik, 1988; Salthouse, Rogan, & Prill, 1984). Specifically, in a prospective duration judgment experiment, in which participants know that they must provide a duration judgment at the end of each trial, more difficult concurrent tasks will permit less attention to the passage of time and will result in shorter verbal estimates of the trial's duration. Given the greater complexity experienced by older adults in dividing attention between tasks (in this case, between the ongoing task and

the judgment of duration), this tendency should be amplified in older participants. Thus, both biological and cognitive perspectives appear to predict that aging should be associated with a decrease in the ratio of subjective to objective duration.

It is therefore surprising to find that the available evidence points in exactly the other direction. In a recent meta-analysis of the literature on age-related differences in human duration judgment, Block, Zakay, and Hancock (1998) found that duration judgment ratios were significantly larger in older adults than in younger adults. That is, the subjective experience of duration was longer in the older group, suggesting either that the internal clock speeds up with age or that more attentional resources are available to older participants. The meta-analysis was based on 14 studies, and it yielded a value of $d_+ = 0.40$, $p < .0001$, which is a moderately strong effect. However, virtually all the studies involved the judgment of an empty interval; typically, participants were asked to produce a given interval (e.g., 30 sec or 3 min) by holding down a key or otherwise indicating the beginning and end of the interval. Only one study is an exception: Polyukhov (1989) reported the typical age-related increase in duration judgment ratio when participants performed no task, but no significant age-related effect when they performed a verbal task. One purpose of the present study was therefore to compare duration judgment ratios of younger and older adults when they were actively performing a demanding perceptual task.

The experiment reported here investigated age-related differences in subjective time estimates using intervals of 30, 60, and 120 sec and employing the methods of verbal estimation ("how long were you working on the task you have just finished?") and production ("continue working on this task until x seconds have elapsed"). Participants knew that time was being measured, so the prospective paradigm was used throughout. In all cases, the interval was filled with a perceptual judgment task that varied in complexity from trial to trial. One well-established finding in the time duration judgment literature is that verbal estimates of duration decrease as task complexity increases (Brown, 1985; Hicks et al., 1976; Smith, 1969; Zakay, 1993; Zakay, Nitzan, & Glicksohn, 1983), possibly because greater attentional demands of the task result in less attention being paid to the passage of time itself (Block & Zakay, 1996; Frankenhaeuser, 1959; Zakay & Block, 1996). Given that, in the present experiment, the intervals were all filled, and, thus, in a sense, the duration judgment task was a secondary task, it was predicted that older participants would report shorter verbal estimates but produce longer intervals in production trials; both cases would reflect the reduced attentional resources available to older participants in dual-task situations. Furthermore, if older adults are penalized relatively more than their younger counterparts by increasing task complexity under dual-task conditions (McDowd & Craik, 1988; Salthouse et al., 1984), it follows that the predicted age-related drop in duration judgment

ment ratios should increase as the complexity of the primary task increases.

A final purpose of the present experiment was to compare the methods of verbal estimation and production. There is some agreement in the literature on duration judgments that different methods may yield different results and may even involve different processes (Allan, 1979; Zakay, 1993). This argument is not very convincing in the cases of prospective verbal estimation and production, since it seems that, in both cases, the participant should check his or her internal clock or cognitive counter and then either report the reading (in verbal estimation) or decide to stop or continue performing the ongoing task (in production). If a person's internal clock is running slow (or little attention is devoted to the passage of time), then when 60 sec of external time has passed, the counter may read 45 sec, which is then reported in the case of verbal estimation. In the case of a 60-sec production task, the participant will continue the task (since the counter reads only 45 sec) and will stop the task when the counter reads 60 sec, at which point 80 sec of real time has elapsed. In both cases, the duration judgment ratios will be 0.75, since, in the verbal estimation case, the subjective duration is 45 sec, and the objective duration is 60 sec; in the case of production, the subjective duration is 60 sec, and the objective duration is 80 sec. If these arguments are correct, the methods of verbal estimation and production should yield equivalent duration judgment ratios.

In summary, in the present study, we investigated age-related differences in judgments of duration by comparing the performance of younger and older adults on time estimation and time production. Intervals of 30, 60, and 120 sec were judged while participants made perceptual judgments at five levels of complexity. The use of three different intervals allowed us to determine the function relating estimated duration to actual duration and thereby to estimate the rate of the internal clock (Salhouse, Wright, & Ellis, 1979). Given that subjective time appears to pass more quickly for older adults, we hypothesized that older participants have fewer attentional resources to devote to the passage of time or that the process of aging is associated with a slowing of the internal clock. Older adults should therefore give shorter verbal estimation values and longer production values than younger adults. The design also enabled us to investigate age-related effects of increasing task complexity on time judgments and to examine the relations between the methods of verbal estimation and production.

METHOD

Participants

In all, 60 people participated in the experiment. Of these, 30 were undergraduates between 18 and 32 years of age (mean age = 22.2 years) with an average of 15.1 years of education; 30 were older adults between 63 and 83 years of age (mean age = 72.2 years) with an average of 13.8 years of education. The younger adults participated for course credit or for a payment of \$8. The older adults were drawn from a pool of volunteers who live at home and participate

in experiments for no payment, although their expenses are reimbursed. The older participants traveled to the lab to be tested. All participants were in good health.

Design and Materials

In overview, all participants performed a visual judgment task at each of five levels of complexity. Each task was performed for 30, 60, and 120 sec under verbal estimation and time production instructions. The design was therefore a 2 (age) \times 5 (task complexity) \times 3 (time interval) \times 2 (type of time judgment) factorial, with age as a between-subjects factor and the other variables as within-subject factors.

The visual task was an absolute judgment task in which stimuli were displayed serially on a computer monitor. Each stimulus was a bright patch that could vary in shape between a circle and a series of progressively flatter ovals; in all, seven different shapes were used, ranging systematically in value from 1 (circle) to 7 (the flattest oval). One task was to judge the shape of a presented stimulus by giving its judged value, 1–7. In addition to changes in shape, each presented stimulus could vary in size, in brightness, and in horizontal and vertical positions on the screen; in all cases, there were seven possible values in each dimension. Thus, for example, a specific stimulus might be located at the 2nd horizontal position and the 5th vertical position, with a shape value of 6, a size value of 1, and a brightness value of 7. The actual values for each stimulus were determined randomly. Task complexity was manipulated by asking the participants to judge only 1 dimension (e.g., brightness, and ignore changes in the other dimensions), 2 dimensions concurrently (e.g., size and horizontal position), or 3, 4, or all 5 dimensions concurrently. Thus, in the 5-dimension task, the participants were asked to judge all 5 dimensions by giving best-guess values for each dimension in turn. In all conditions, stimuli were displayed for the same length of time (10 sec); thus, judgments of only 1 dimension made the task easy and somewhat boring, judgments of 3 specified dimensions made the task more difficult, and judgments of all 5 dimensions constituted a very demanding task. Before each trial, the participant was told which dimensions were relevant (e.g., shape, horizontal, brightness), and these same dimensions remained relevant throughout the trial. To minimize confusion, as each stimulus appeared, the experimenter called out the first dimension to be judged (e.g., "shape") and then the second and subsequent dimensions after the participant responded by saying the judged value. The experimenter recorded the participants' responses, thereby minimizing the possibility that they could infer the passage of time by looking at the number of stimuli they had judged.

As well as performing the absolute judgment task, the participants made verbal estimates (or productions) with respect to the duration of the current trial. Half of the trials involved verbal estimation, and half involved production. In the case of verbal estimation, the participants were instructed to keep working on each trial until the computer flashed the message "How long have you been working on this trial?" The participants were told to type their responses into the computer after this message appeared. They were also told that the durations of trials were not obvious lengths of time and that they should therefore not give estimates such as "1 minute" but rather give estimates to the nearest judged second. In the case of time production, the participants were given an interval in seconds before each trial commenced. This interval was written on a cue card and placed in front of the participants to avoid forgetting of the current time production task. The participants were also told to keep working at the visual estimation task until they judged that the specified duration had elapsed, at which point they should terminate the trial by pressing any key on the computer keyboard. The computer recorded each produced duration. It was emphasized that they should stop the trial as soon as the specified interval was reached, regardless of whether they had finished rating the current stimulus. Presentation of verbal estimation and production trials was blocked; the

participants performed 15 estimation or production trials initially and only then received instructions for the other condition. The order in which the participants performed the blocks was randomly determined. Within each block, the participants performed the combination of three time intervals (30, 60, 120 sec) and five levels of task complexity; the order of these combinations and the specific dimensions chosen for each trial were randomly determined for each participant separately. A final feature of the duration judgment task was that the actual intervals presented included intervals that were 3 and 6 sec greater and less than the nominal target duration. Thus, for example, the five 60-sec intervals used were 54, 57, 60, 63, and 66 sec. This reduced the participants' ability to guess that the same three standard intervals were being used throughout and also served to emphasize the point that we did not wish the participants to round up or down to more standard intervals.

Procedure

The participants were tested individually. They were first given practice on the visual judgment task. After the general principles were explained, they judged 14 different stimuli in which all seven values were presented twice for each dimension. The participants made judgments on all 5 dimensions for each stimulus, and they received feedback on the accuracy of each response. On completion of the practice session, the participants were informed that they would be carrying out a duration judgment task while performing the visual task. Depending on the order of blocks, they then received instructions for either the duration estimation task or the duration production task. Before starting each block of judgments, the participants were given a further 75-sec practice trial, rating all 5 dimensions. The stimulus presentations in this practice trial were presented at the same rate as in the actual test—that is, each stimulus was displayed for 9.5 sec, preceded by a 0.5-sec, "ready" message. The participants were not informed that each stimulus was displayed for 9.5 sec, however. Before each test trial, the participants were told which dimensions they would be judging, and they were given the opportunity to review the full range of values on each relevant dimension at their own pace. The participants were not told how much time they would have to make their judgments, but they were told that, when judging only one dimension, the task would be easy and that they would have plenty of time to give their answer; however, when judging all five dimensions, they were told that the task would be difficult and that they would have to work quickly to finish in time. If they did not finish their judgments before the next stimulus appeared, they were told to continue with the new stimulus. It was stressed, however, that it was more important to finish rating all the dimensions required of them than it was to be extremely accurate on some of the dimensions but not to complete them all.

RESULTS

Visual Judgment Task

Performance on the visual judgment task was measured in terms of information transmission. The task involved absolute judgments of 1–5 dimensions presented simultaneously, and the participants' performance on each dimension was assessed by calculating the relation between input and output variables in the manner suggested by Attneave (1959). Since there were always seven possible alternative values for each stimulus on each dimension, input information was $\log_2 7 = 2.81$ bits per stimulus dimension, and total input information therefore ranged from 2.81 to 14.05 bits per stimulus on each trial, depending on how many of the 5 dimensions were involved. Transmitted information was calculated for each

Table 1
Transmitted Information (Bits Per Stimulus and Mean Bits Per Dimension) as a Function of Age and Number of Judged Dimensions

Information	Dimensions				
	1	2	3	4	5
Young Age Group					
Total bits per stimulus	1.60	3.09	4.73	6.04	7.31
Mean bits per dimension	1.60	1.55	1.58	1.51	1.46
Old Age Group					
Total bits per stimulus	1.52	3.00	4.34	5.32	5.50
Mean bits per dimension	1.52	1.50	1.45	1.33	1.10

Note—The values shown are averaged over duration judgment method and time interval.

dimension and then summed to provide a measure of performance. These values are shown in Table 1, both as total bits per stimulus and as mean bits transmitted on each dimension.

The table shows that, for both younger and older participants, the number of bits per stimulus transmitted increased substantially as the number of dimensions to be judged increased from 1 to 5. These results show unequivocally that the number of successful (or partially successful) discriminations increased as more dimensions were added. This increase was somewhat less for the older participants, as would be expected if processing capacity declines with age (Welford, 1958). The question of changes in subjective difficulty was not answered directly by these data, but it was addressed to some extent by the changes in mean bits per dimension transmitted as a function of number of dimensions processed. These latter values declined in general from 1 to 5 dimensions for both age groups, showing that performance on any one dimension was best when it was judged alone and was worst (presumably most difficult) when it was judged along with 4 other dimensions. On the basis of the data shown in Table 1, we argue that our manipulation of task complexity was effective: The participants had to perform more discriminative work as the number of judged dimensions was increased from one to five.

Duration Judgments

The means for verbal estimations and productions are shown graphically in Figure 1. The figure shows that, in all cases, the verbal estimates were under the actual times, whereas the productions were over; it also shows that this pattern was greatly exaggerated in the older participants. There was surprisingly little evidence of an effect of task complexity, except in the case of the 120-sec judgments. The expected result was that verbal estimates would decline and that productions would lengthen as complexity increased with the number of dimensions to be judged. The effects of age and task complexity were assessed by a series of analyses of variance (ANOVAs). For verbal estimations, the analysis for 30 sec showed a significant effect for age group [$F(1,58) = 11.99, MS_e = 452.7, p <$

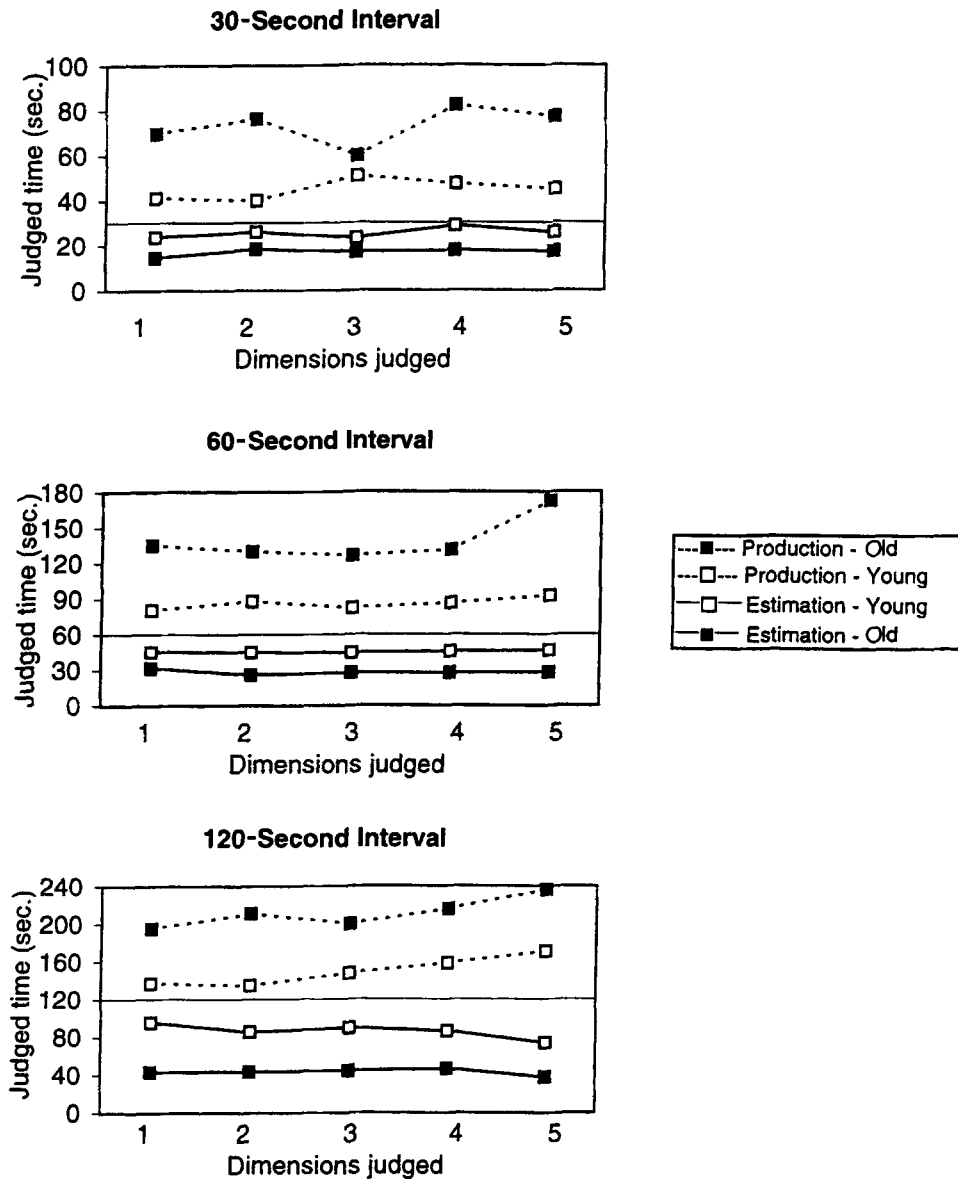


Figure 1. Duration production and verbal estimation values for younger and older adults as a function of task interval and number of dimensions judged.

.01] and an effect of number of dimensions judged [$F(4,232) = 2.62$, $MS_e = 58.9$, $p < .05$], but no reliable interaction. Older adults gave lower verbal estimates than younger adults, and, unexpectedly, estimates increased slightly, if anything, as more dimensions were added. The ANOVA for 60 sec (verbal estimation) revealed a significant effect of age [$F(1,58) = 12.71$, $MS_e = 1,742$, $p < .001$], but no effect of dimensions or of the interaction. Again, older participants gave smaller estimates. The analysis for 120 sec (verbal estimation) showed an effect of age [$F(1,58) = 23.35$, $MS_e = 6,032$, $p < .001$] and a significant effect of number of dimensions [$F(4,232) = 2.83$, $MS_e = 590.6$, $p < .05$], but no significant interaction. Older participants gave smaller estimates, and more

complex tasks (i.e., more dimensions to be judged) were associated with shorter verbal estimates. For the production data, the analysis for 30 sec showed an effect of age [$F(1,58) = 16.47$, $MS_e = 3,617$, $p < .01$], but no reliable effects of dimensions or of the interaction. In this case, older participants produced longer time intervals. The ANOVA for 60 sec (production) revealed that the effect of age was reliable [$F(1,58) = 29.52$, $MS_e = 7,153$, $p < .001$], as was the effect of dimensions [$F(4,232) = 3.16$, $p < .02$]. The former effect was again attributable to older participants' producing longer intervals, and the latter effect was attributable to the higher production values for the 5-dimension condition. The interaction was not significant. Finally, the ANOVA for 120 sec (pro-

duction) showed a significant effect of age [$F(1,58) = 25.96$, $MS_e = 11,002$, $p < .001$] and a significant effect of dimensions [$F(4,232) = 5.53$, $MS_e = 2,339$, $p < .001$]. The interaction was not significant. Figure 1 shows that longer produced times were associated with older participants and more judged dimensions.

A second way to view the data is to plot duration estimates and productions against the three time intervals used in the experiment (30, 60, 120 sec). The values of this method are (1) to see if the functions are linear, as they should be if judged duration is some constant ratio of actual duration, (2) to find these ratios for the two age groups, and (3) to determine whether duration judgment ratios for production approximate the ratios for verbal estimation. The third possibility follows from the assumption that participants make their judgments on the basis of the same hypothesized internal clock or counter for verbal estimation and production. Figure 2 shows the best-fit linear functions for verbal estimation and production (collapsed over the five levels of task difficulty) for each age group separately. The figure shows that the mean duration judgments were well fitted graphically by linear functions; regression equations of actual duration on judged duration gave F values for the linear contrast of 225.4, 305.2, 119.9, and 29.6 for old-production, young-production, young-estimation, and old-estimation, respectively ($df = 1,29$, $p < .001$, in all cases).

Table 2
Slope Constants for Best-Fit Linear Functions
Relating Judged Duration to Actual Duration
as a Function of Age and Experimental Condition

Perceptual Task	Verbal Estimation		Production	
	Young	Old	Young	Old
1 Dimension	0.81	0.30	1.05	1.34
2 Dimensions	0.67	0.28	1.02	1.48
3 Dimensions	0.74	0.30	1.08	1.52
4 Dimensions	0.64	0.32	1.23	1.47
5 Dimensions	0.52	0.22	1.39	1.67

Table 2 shows the slope constants of the best-fit linear functions relating judged duration to actual duration when the data were broken down by task complexity (i.e., by the number of dimensions judged in the absolute judgment task). In the case of verbal estimation, Table 2 shows that the slope constants were consistently greater for younger participants than for older participants and that all values of the slope constant were less than 1.0. Following the logic laid out in the introduction, these results indicate either that the internal clock was running slow in all cases in this experiment and that aging is associated with a greater degree of slowing or that older participants devoted less attention to the passage of time. Averaging over the two age groups, the slope constants also showed some tendency to decrease in value as the number of

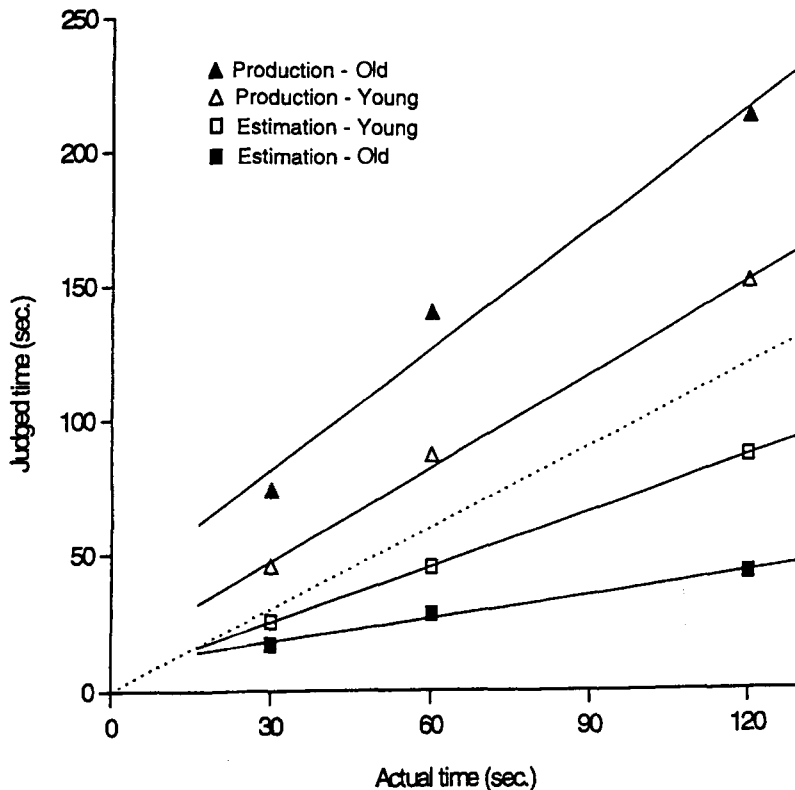


Figure 2. Best-fit linear functions relating judged duration to actual duration as a function of age and method.

judged dimensions increased ($M_s = 0.56, 0.48, 0.52, 0.48, \text{ and } 0.37$, for 1, 2, 3, 4, and 5 dimensions, respectively). That is, there was a tendency for more complex (and possibly more difficult) tasks to be associated with greater degrees of underestimation. In the case of production, Table 2 shows that all slope values were greater than 1.0, that the values were all greater for older participants than for younger participants, and that, averaging over age groups, there was a strong trend for slope values to increase with number of judged dimensions ($M_s = 1.20, 1.25, 1.30, 1.35, \text{ and } 1.53$ for 1, 2, 3, 4, and 5 dimensions, respectively). That is, more complex tasks were associated with greater degrees of overproduction.

Relations Between Verbal Estimation and Production

As described previously, if duration judgments using the methods of verbal estimation and production reflect the same cognitive processes and/or biological mechanisms, the ratios of subjective to actual durations should be equivalent for the two methods. Also, the corresponding ratios for the two methods should be affected similarly by other experimental variables (in this case, age, task complexity, and task duration). Figure 3 shows the observed duration judgment ratios as a function of these variables. Note that the ratios for the method of production were all less than 1.0, despite the finding that the

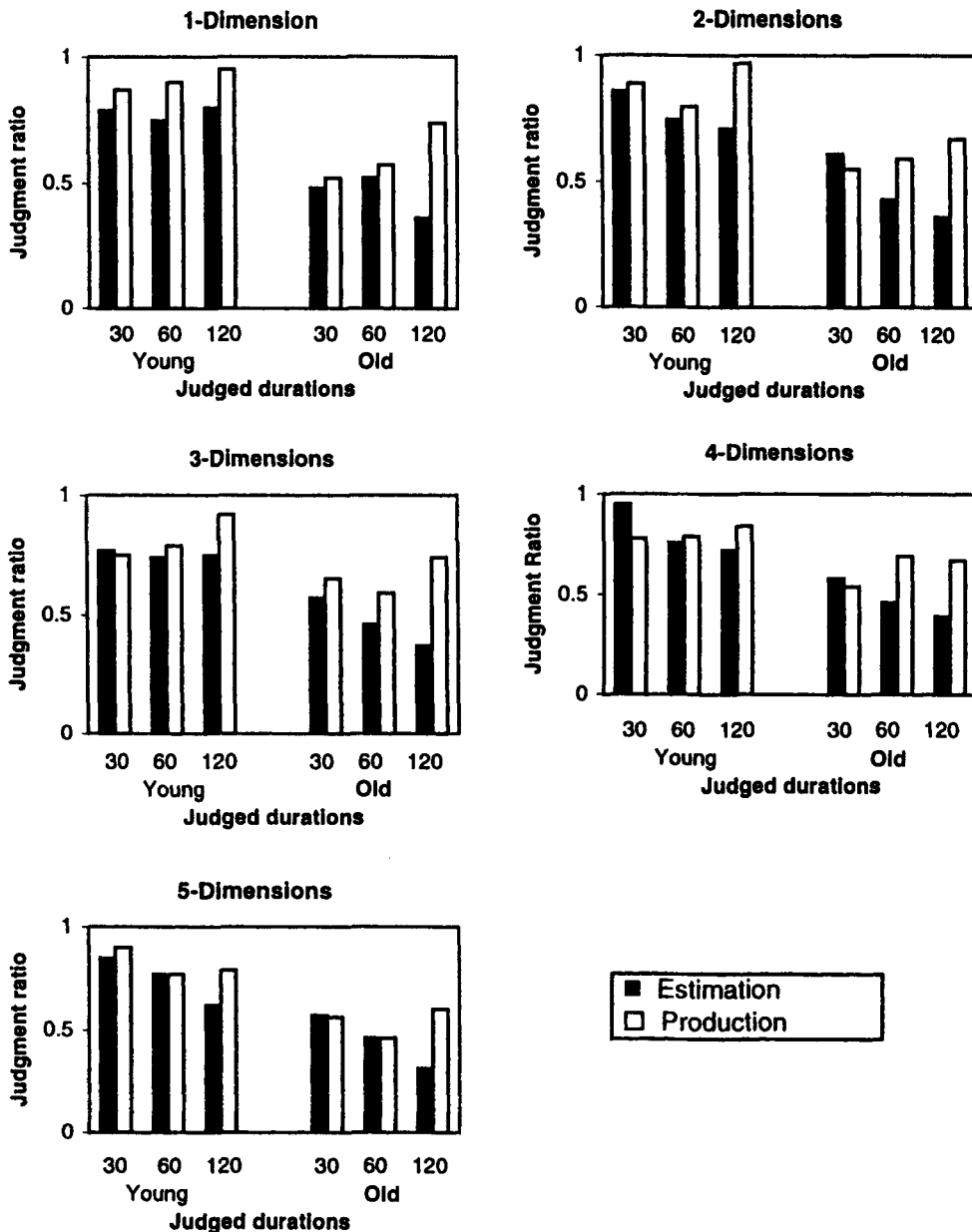


Figure 3. Duration judgment ratios as a function of task complexity, task duration, age group, and method.

participants produced intervals longer than the target intervals (Figure 1). This occurred because the duration judgment ratios shown in Figure 3 were all calculated as the ratio of subjective to objective duration, and, in the case of production, the subjective duration was the specified target duration (e.g., 120 sec), which was typically shorter than the objective (produced) duration.

Differences between the ratios as a function of type of judgment (verbal estimation vs. production), age group, task duration (30, 60, 120 sec), and task complexity (1–5 dimensions) were assessed by means of a four-way ANOVA in which age was a between-subject factor, and judgment, duration, and complexity were within-subjects factors. The analysis yielded main effects of age [$F(1,58) = 24.6$, $MS_e = 1.377$, $p < .001$], type of judgment [$F(1,58) = 7.98$, $MS_e = 0.694$, $p < .01$], and duration [$F(2,116) = 3.53$, $MS_e = 0.119$, $p < .05$]. Overall values of the ratios were greater for young participants than for older participants (0.81 and 0.54, respectively), greater for production than for verbal estimation (0.73 and 0.62, respectively), and somewhat greater for shorter durations (0.70, 0.65, and 0.66 for 30, 60, and 120 sec, respectively). In addition, two interactions were statistically reliable: type of judgment \times task duration [$F(2,116) = 34.11$, $MS_e = 2.511$, $p < .001$] and task complexity \times task duration [$F(8,464) = 3.62$, $MS_e = 0.180$, $p < .001$]. The first interaction reflects the fact that verbal estimation ratios decreased with increasing durations (0.70, 0.61, and 0.54 for 30, 60, and 120 sec, respectively), whereas production ratios tended to increase with increasing durations (0.70, 0.70, and 0.79 for 30, 60, and 120 sec, respectively). Values for the task complexity \times task duration interaction are shown in Table 3. The table shows that the ratios increased slightly from 1 to 5 dimensions for 30 sec, decreased slightly for 60 sec, and decreased markedly for 120 sec.

These comparisons between the two types of time judgments yielded a complex but understandable picture. First, the significant age difference again reflects the finding that older participants were more extreme in their judgments, showing both greater degrees of underestimation in the verbal estimation task and greater degrees of overproduction in the production task. The finding that production ratios were larger than verbal estimation ratios was probably an artifact stemming from the use of the participants' individual ratios in the case of produc-

tion scores.¹ When the corresponding scores were calculated from the overall means of production and estimation scores, the values were 0.60 and 0.62, respectively.

The small effect of task duration was qualified by a large interaction between duration and type of judgment: The proportions tended to drop from 30 sec to 120 sec in the case of verbal estimation, but rose from 30 sec to 120 sec in the case of production (see Figure 3). This pattern of results probably reflects the "series effect," or central tendency, in time judgments (see Woodrow, 1951; Woodworth, 1938), whereby participants tend to overestimate short intervals within the experimental range but underestimate long intervals. In the present experiment, all subjective duration judgments tended to be less than the actual durations, owing to other factors (e.g., age and task difficulty), but Figure 3 shows that, in the case of verbal estimation, the underestimations were least for 30 sec and greatest for 120 sec. In the case of duration productions, however, the central tendency effect would act to raise productions of shorter intervals but lower productions of longer intervals, resulting in relatively lower judgment ratios for 30-sec intervals and relatively higher ratios for 120-sec intervals. These effects are also seen in Figure 3. The final significant effect from the ANOVA was the interaction between task duration and task complexity shown in Table 3. To be in line with previous findings, ratios should decrease from 1 to 5 dimensions (i.e., verbal estimations decrease and productions increase as task complexity increases). This effect was present for 120 sec but was absent for 30 sec. In retrospect, the shorter intervals might have given the participants too many cues to passing time, given that each stimulus was exposed for a constant 10-sec interval.

These findings have several implications for the interpretation of estimation and production ratios as measures of the rate of accumulation of pulses in the temporal counter (or rate of the "biological clock"). Most obviously, the ratios varied as a function of task duration, arguably due to central tendency or range effects. Thus, the average ratios may give a reasonable estimate of the counter or clock's rate in each condition, whereas the ratios for durations at the extremes of the range clearly do not. This point is also relevant to the postulated equivalence of duration judgment ratios calculated from the methods of verbal estimation and production. That is, this equivalence should be found only for ratios averaged across the range of time intervals used. In the present experiment, the average verbal estimation ratio for young participants is 0.77, and the corresponding production ratio was 0.72 (calculated from the group means, not from the means of individual production ratio scores; see note 1). For older participants, the corresponding ratios were 0.46 and 0.47, respectively. Overall, then, the ratios were 0.62 and 0.60 for verbal estimation and production, respectively; the proximity of these values suggests that both methods reflect a common time perception mechanism. However, this comparison should be treated cautiously given that the

Table 3
Mean Duration Judgment Ratios as a
Function of Task Duration and Task Complexity

Task Complexity	Task Duration (sec)		
	30	60	120
1 Dimension	0.67	0.69	0.71
2 Dimensions	0.73	0.64	0.68
3 Dimensions	0.69	0.65	0.70
4 Dimensions	0.71	0.68	0.66
5 Dimensions	0.72	0.62	0.58

ANOVA based on each participant's production scores yielded means of 0.62 and 0.73 for verbal estimation and production, respectively—a significant difference.

In summary, the present experiment showed that older people give shorter verbal duration estimates and longer duration production scores than do their younger counterparts. In line with previous studies, increased task complexity decreased verbal estimates and increased productions, although only with the 120-sec interval. In accordance with the notion that verbal estimation and production methods reflect a common mechanism, the manipulations of age group and task complexity affected the two methods similarly. Direct comparisons of verbal estimation ratios and production ratios were complicated by range effects and by the (possibly) inflated estimates of production ratios when these scores were obtained by averaging over the participants' individual ratios.

DISCUSSION

The present experiment was carried out primarily to examine possible age-related differences in judgments of duration, but also with the secondary aims of comparing the methods of verbal estimation and production and of gathering further information on the effects of different tasks on duration judgments. Our predictions were that older participants would give shorter verbal estimations but longer productions than would younger participants (implying a slower internal clock or less attention paid to the passage of time in the older group), that duration judgment ratios for two methods would be approximately equivalent, and that more demanding tasks would be associated with shorter verbal estimations but longer time productions. These predictions are now discussed in relation to the results obtained.

Task Effects

Task effects were generally small in the present study; the participants might have based their responses on the number of trials performed (all trials were 10 sec long), and this might have overridden task effects at the shorter intervals. In the 120-sec interval conditions, task complexity did have an effect: The more complex conditions were associated with shorter verbal estimations and longer productions, in line with previous findings (Block, Conti, Duncan, & Zakay, 1997).

Method Effects

If duration judgments are based on the same set of cognitive and biological processes, then estimates derived from different methods should yield the same approximate values. This prediction was tested in the present experiment by comparing duration judgment ratios (the ratios of subjective to actual durations) for the methods of verbal estimation and duration production. The comparisons were complicated by two factors: the presence of strong range effects, and the fact that mean production ratios can be biased by a few anomalous judgments. The

Table 4
Duration Judgment Ratios for the Main Conditions Calculated From Group Means

Condition	Time Interval (sec)			<i>M</i>
	30	60	120	
Young Age Group				
Estimation	0.84	0.76	0.72	0.77
Production	0.67	0.70	0.81	0.72
Old Age Group				
Estimation	0.56	0.47	0.36	0.46
Production	0.41	0.44	0.57	0.47

Note—The values given are averaged over tasks.

effects of the latter factor can be minimized by calculating duration judgment ratios from group means rather than from individual scores before averaging. Table 4 shows the ratios calculated in this way, averaged over tasks. The table reveals the strong tendency for verbal estimation ratios to decrease as the time interval lengthens but for production ratios to increase over the same range of intervals.² As discussed previously, this pattern stems from the tendency of the participants to give relatively long verbal estimates and long productions for the shortest interval in the range used and relatively short verbal estimates and productions for the longest intervals used. In both cases, the range of subjective judgments is truncated relative to the actual range. The presence of range effects means that it is not sensible to compare individual values of duration judgment ratios, but it should be possible to compare mean values averaged across each range of time intervals. These means are also shown in Table 4. The resulting mean values were reasonably close for the two methods; when these means were averaged over the younger and the older participants, the overall averages for verbal estimation and production were 0.62 and 0.60, respectively. We therefore conclude that the two methods do reflect the same underlying processes and mechanisms.

A final implication of the present results is that although, at first sight, it seems reasonable to estimate the rate of some postulated internal clock from the function relating duration judgments to actual durations (e.g., Salthouse et al., 1979), the likely presence of range effects reduces this method's validity. If the subjective range is truncated, as in the present experiments, the resulting slope will be flattened and the rate underestimated. From the present results, it seems that a more valid index of the rate of subjective time can be obtained from the mean value of the duration judgment ratios averaged across the time intervals used.

Age Effects

The present experiment showed large and robust age-related effects on judgments of duration. Older participants exhibited the same pattern of results with respect to tasks and intervals as their younger counterparts but gave consistently shorter estimates in the verbal estima-

tion trials and produced consistently longer intervals in the production trials. These results are in line with the observation that external clock time appears to pass faster for older people—that is, when they estimate that 30 sec has passed, 60 sec of clock time has actually elapsed, and when asked to produce 60 sec, they produce 120 sec. The present findings are thus in agreement both with biologically oriented models that postulate an age-related slowing of some internal clock (see Schroots & Birren, 1990, for discussion) and with cognitive models that postulate duration judgment processes that depend on the amount of attention paid to time passing (Block & Zakay, 1996; Zakay & Block, 1996). In the latter case, the suggestion is that older people are more penalized by the need to divide attention between some ongoing task and the temporal judgment and so devote less attention to the passage of time.

It was also predicted that age would interact with task complexity, in the sense that, as the perceptual task became more difficult, older participants would show differentially shorter verbal estimates and differentially longer productions. The assessment of this prediction was complicated by the relatively small effects of task complexity found in the present experiment; however, for the 120-sec interval condition, in which complexity had quite large effects, there was no trace of the predicted interaction. Increasing task complexity was associated with shorter estimates and longer productions, but to the same extent for younger and older participants.

Do these results address the question of the relative importance of biological and cognitive factors in time perception? The preceding discussion suggests that if the smaller duration judgment ratios observed in older adults were attributable to reduced attentional resources, the age-related differences in the ratios should have increased with increasing task complexity, yet no such effect was found. The finding of a large main effect of aging but no interaction with task complexity thus favors a slowing of the hypothesized clock or pacemaker, rather than a reduction in attentional resources, as the underlying cause of the age-related changes in judgments of duration. This conclusion is drawn quite tentatively, however, given that the effect of task complexity was small in the present experiment. It also seems likely that people would adapt to any long-term permanent drift in the rate of a pacemaker, in the same way that they adapt to distortions in visual input caused by prisms or lenses (e.g., Held, 1965; Rock & Harris, 1967). Strong effects of aging on time perception may appear only under unfamiliar conditions and with an atypically high workload, as in the present experiment.

Recent findings from neuropsychological studies of patients with frontal lobe lesions give some support to this speculation. Casini and Ivry (1999) and Mangels, Ivry, and Shimizu (1998) found that patients with unilateral lesions in the dorsolateral prefrontal cortex (predominantly in the left hemisphere) were impaired in a judgment discrimination task and that their performance was poorer under dual-task conditions (Casini & Ivry, 1999). Binkofski and Block (1996) reported the case of

a patient with a tumor in his left superior prefrontal cortex; when asked to produce durations of 60 sec, his mean production was 286 sec, giving a subjective duration ratio of 0.24. Although the neuropsychological evidence is clearly at an early stage, it is nonetheless interesting given the parallels that some researchers have drawn between the effects of frontal lobe lesions and the effects of normal aging on cognitive processing (Craik, Morris, Morris, & Loewen, 1990; West, 1996).

The remaining puzzle is why the present experiment showed such clear effects of aging that are in direct opposition to the findings of the recent meta-analysis (Block et al., 1998). The most likely solution to this puzzle is that the present experiment involved participants in a demanding perceptual task, whereas all 14 studies included in the meta-analysis used empty time intervals. One of these studies (Polyukhov, 1989) also included judgments of intervals filled with a sentence creation task. The researcher found that older participants gave higher verbal estimates of the 45-sec interval than did their younger counterparts in the empty interval case, whereas there were no significant age differences in the case of filled intervals. Additionally, and in line with the present experiment, participants of all ages underestimated the interval after performing the sentence creation task. The results of a study by Mangan, Bolinsky, and Rutherford (1997) (not included in the Block et al., 1998, meta-analysis) are also in line with the present findings. These researchers found that older adults produced longer durations than did their younger counterparts when asked to judge 60 and 180 sec (i.e., duration judgment ratios were smaller for the older participants), and they also found that this effect was stronger when participants engaged in a demanding cognitive task during the judged interval.

More than half of the studies in the Block et al. (1998) meta-analysis were conducted before 1980 and used a relatively simple methodology. For example, Feifel (1957) asked participants to produce durations of 30, 60, 180, and 300 sec using any method they wished, and this procedure was followed by several other investigators. Under these conditions, many participants simply counted out the required number of seconds. A more recent study using a more sophisticated methodology was reported by Eisler and Eisler (1994). Using the method of reproduction and a psychophysically based clock model, these investigators concluded that older adults show longer subjective durations than do younger adults—that is, a result in line with the meta-analysis of Block et al. (1998) but opposite to the present result. Again, however, Eisler and Eisler used empty time intervals. The Eislers suggest that the apparent contradiction between their findings and the general observation that time seems to pass more quickly with advancing age may be resolved if the everyday observation is based on retrospective judgments of time that has passed, as opposed to judgments of time in progress (as in their experiment)—the prospective paradigm. Contrary to this suggestion, however, in the present experiment, we also used the prospective paradigm

yet found older age to be associated with shorter subjective durations. Our tentative conclusion is therefore that empty intervals are typically overestimated to a greater extent by older adults, whereas intervals filled with cognitive activity are underestimated to a greater extent by older adults.

In conclusion, while the present findings cannot be used unequivocally to disentangle possible long-term biological factors from task- and context-related psychological factors in the perception of time, they do provide some additional empirical evidence on some issues. First, robust age-related effects were documented, with older adults showing a greatly exaggerated tendency to underestimate and overproduce intervals of time under the present experimental conditions. Second, the approximate equivalence of verbal estimation and production methods was illustrated once range effects were taken into account. Finally, the results endorse recent contextually dependent views of time perception (e.g., Allan, 1979; Block, 1990; Zakay, 1990). To understand and interpret a given duration judgment, such factors as task, method, range, and age must be taken into account.

REFERENCES

- ADAM, N., ROSNER, B. S., HOSICK, E. C., & CLARK, D. L. (1971). Effect of anesthetic drugs on time production and alpha rhythm. *Perception & Psychophysics*, **10**, 133-136.
- ALLAN, L. G. (1979). The perception of time. *Perception & Psychophysics*, **26**, 340-354.
- ATTNEAVE, F. (1959). *Applications of information theory to psychology*. New York: Holt.
- BINKOFSKI, F., & BLOCK, R. A. (1996). Accelerated time experience after left frontal cortex lesion. *Neurocase*, **2**, 485-493.
- BLOCK, R. A. (1990). Models of psychological time. In R. A. Block (Ed.), *Cognitive models of psychological time* (pp. 1-35). Hillsdale, NJ: Erlbaum.
- BLOCK, R. A., CONTI, D. E., DUNCAN, A. M., & ZAKAY, D. (1997, November). *Workload influences on duration judgments: Meta-analytic findings*. Poster presented at the annual meeting of the Psychonomic Society, Philadelphia.
- 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.
- BLOCK, R. A., & ZAKAY, D. (1996). Models of psychological time revisited. In H. Helfrich (Ed.), *Time and mind* (pp. 171-195). Kirkland, WA: Hogrefe & Huber.
- BLOCK, R. A., & ZAKAY, D. (1997). Prospective and retrospective duration judgments: A meta-analytic review. *Psychonomic Bulletin & Review*, **4**, 184-197.
- BLOCK, R. A., ZAKAY, D., & HANCOCK, P. A. (1998). Human aging and duration judgments: A meta-analytic review. *Psychology & Aging*, **13**, 584-596.
- BROWN, S. W. (1985). Time perception and attention: The effects of prospective versus retrospective paradigms and task demands on perceived duration. *Perception & Psychophysics*, **38**, 115-124.
- CASINI, L., & IVRY, R. B. (1999). *Effects of divided attention on temporal processing in patients with lesions of the cerebellum or frontal lobe*. Manuscript submitted for publication.
- CHURCH, R. M. (1984). Properties of the internal clock. In J. Gibbon & L. Allan (Eds.), *Timing and time perception* (Annals of the New York Academy of Sciences, Vol. 423, pp. 566-582). New York: New York Academy of Sciences.
- CRAIK, F. I. M. (1977). Age differences in human memory. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (pp. 384-420). New York: Van Nostrand Reinhold.
- CRAIK, F. I. M., & BYRD, M. (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. Trehub (Eds.), *Aging and cognitive processes* (pp. 191-211). New York: Plenum.
- CRAIK, F. I. M., MORRIS, L. W., MORRIS, R. G., & LOEWEN, E. R. (1990). Aging, source amnesia, and frontal lobe functioning. *Psychology & Aging*, **5**, 148-151.
- EISLER, A. D., & EISLER, H. (1994). Subjective time scaling: Influence of age, gender, and type A and type B behavior. *Chronobiologia*, **21**, 185-200.
- FEIFEL, H. (1957). Judgment of time in younger and older persons. *Journal of Gerontology*, **12**, 71-74.
- FRAISSE, P. (1984). Perception and estimation of time. *Annual Review of Psychology*, **35**, 1-36.
- FRANKENHAEUSER, M. (1959). *Estimation of time: An experimental study*. Stockholm: Almqvist & Wiksell.
- FRIEDMAN, W. J. (1990). *About time: Inventing the fourth dimension*. Cambridge, MA: MIT Press.
- GOLDSTONE, S., & KIRKHAM, J. E. (1968). The effects of secobarbital and dextroamphetamine upon time judgment: Intersensory factors. *Psychopharmacologia*, **13**, 65-73.
- HELD, R. (1965, November). Plasticity in sensory-motor systems. *Scientific American*, **213** (5), 86-94.
- HICKS, R. E., GUALTIERI, C. T., MAYO, J. P., JR., & PEREZ-REYES, M. (1984). Cannabis, atropine and temporal information processing. *Neuropsychobiology*, **12**, 229-237.
- 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.
- JONES, M. R., & BOLTZ, M. (1989). Dynamic attending and responses to time. *Psychological Review*, **96**, 459-491.
- JOUBERT, C. E. (1983). Subjective acceleration of time: Death anxiety and sex differences. *Perceptual & Motor Skills*, **57**, 49-50.
- LEMLICH, R. (1975). Subjective acceleration of time with aging. *Perceptual & Motor Skills*, **41**, 235-238.
- MANGAN, P. A., BOLINSKEY, P. K., & RUTHERFORD, A. L. (1997). Underestimation of time during aging: The result of age-related dopaminergic changes? *Society for Neuroscience Abstracts*, **23**, 2003.
- MANGELS, J. A., IVRY, R. B., & SHIMIZU, N. (1998). Dissociable contributions of the prefrontal and neocerebellar cortex to time perception. *Cognitive Brain Research*, **7**, 15-39.
- MCDOWD, J. M., & CRAIK, F. I. M. (1988). Effects of aging and task difficulty on divided attention performance. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 267-280.
- MECK, W. H. (1996). Neuropharmacology of timing and time perception. *Cognitive Brain Research*, **3**, 227-242.
- MICHON, J. A. (1990). Implicit and explicit representations of time. In R. A. Block (Ed.), *Cognitive models of psychological time* (pp. 37-58). Hillsdale, NJ: Erlbaum.
- ORNSTEIN, R. E. (1969). *On the experience of time*. Harmondsworth, U.K.: Penguin.
- POLYUKHOV, A. M. (1989). Subjective time estimation in relation to age, health, and interhemispheric brain asymmetry. *Zeitschrift für Gerontologie*, **22**, 79-84.
- POYNTER, W. D. (1983). Duration judgment and the segmentation of experience. *Memory & Cognition*, **11**, 77-82.
- ROCK, I., & HARRIS, C. S. (1967, May). Vision and touch. *Scientific American*, **216** (5), 96-104.
- SALTHOUSE, T. A., ROGAN, J. D., & PRILL, K. A. (1984). Division of attention: Age differences on a visually presented memory task. *Memory & Cognition*, **12**, 613-620.
- SALTHOUSE, T. A., WRIGHT, R., & ELLIS, C. L. (1979). Adult age and the rate of an internal clock. *Journal of Gerontology*, **34**, 53-57.
- SCHROOTS, J. J. F., & BIRREN, J. E. (1990). Concepts of time and aging in science. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of the psychology of aging* (3rd ed.). San Diego: Academic Press.
- SMITH, N. C. (1969). The effect on time estimation of increasing the complexity of a cognitive task. *Journal of General Psychology*, **81**, 231-235.

- TINKLENBERG, J. R., ROTH, W. T., & KOPPELL, B. S. (1976). Marijuana and ethanol: Differential effects on time perception, heart rate, and subjective response. *Psychopharmacology*, **49**, 275-279.
- TREISMAN, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock." *Psychological Monographs: General & Applied*, **77** (No.13, Whole No. 576), 1-13.
- WALKER, J. L. (1977). Time estimation and total subjective time. *Perceptual & Motor Skills*, **44**, 527-532.
- WELFORD, A. T. (1958). *Ageing and human skill*. London: Oxford University Press.
- WEST, R. L. (1996). An application of prefrontal cortex function theory to aging. *Psychological Bulletin*, **120**, 272-292.
- WOODROW, H. (1951). Time perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1224-1236). New York: Wiley.
- WOODWORTH, R. S. (1938). *Experimental psychology*. New York: Holt.
- ZAKAY, D. (1990). The evasive art of subjective time measurement: Some methodological dilemmas. In R. A. Block (Ed.), *Cognitive models of psychological time* (pp. 59-84). Hillsdale, NJ: Erlbaum.
- ZAKAY, D. (1993). Time estimation methods: Do they influence prospective duration estimates? *Perception*, **22**, 91-101.
- ZAKAY, D., & BLOCK, R. A. (1996). The role of attention in time estimation processes. In M. A. Pastor & J. Artieda (Eds.), *Time, internal clocks and movement* (pp. 143-164). Amsterdam: Elsevier.
- 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 mismatch stems from the averaging of ratios with different denominators. For example, if 4 participants were asked to produce an interval of 100 sec and produced values of 120, 100, 90, and 50 sec, their individual duration judgment ratios would be 0.83, 1.00, 1.11, and 2.00, respectively, yielding an average of 1.24. However, the group mean production value is 90 sec, giving a duration judgment ratio of 1.11.

2. We carried out a second experiment on 16 young adults to confirm this observation. In this experiment, participants gave verbal estimates and productions for durations of 60, 120, and 180 sec, while performing continuous cognitive tasks. Mean duration judgment ratios for verbal estimation were 0.92, 0.87, and 0.84 for these three durations, respectively; the corresponding ratios for production were 0.78, 0.95, and 1.04. The interaction between type of judgment and task duration was again reliable [$F(2,30) = 6.89$, $MS_e = 0.09$, $p < .01$].

(Manuscript received October 31, 1997;
revision accepted for publication March 20, 1998.)