Journal of Experimental Psychology:
Human Perception and Performance
2007, Vol. 33, No. 3, 716–729

0096-1523/07/$12.00 DOI: 10.1037/0096-1523.33.3.716

Internal Clock Processes and the Filled-Duration Illusion

John H. Wearden
Keele University

Roger Norton, Simon Martin, and
Oliver Montford-Bebb
Manchester University

In 3 experiments, the authors compared duration judgments of filled stimuli (tones) with unfilled ones (intervals defined by clicks or gaps in tones). Temporal generalization procedures (Experiment 1) and verbal estimation procedures (Experiments 2 and 3) all showed that subjective durations of the tones were considerably longer than those of unfilled intervals defined either by clicks or gaps, with the unfilled intervals being judged as approximately 55%–65% of the duration of the filled ones when real duration was the same. Analyses derived from the pacemaker–switch–accumulator clock model incorporated into scalar timing theory suggested that the filled/unfilled difference in mean estimates was due to higher pacemaker speed in the former case, although conclusively ruling out alternative interpretations in terms of attention remains difficult.

Keywords: time perception, temporal generalization, verbal estimation, filled duration illusion, click trains

It has long been known that a variety of factors, in addition to actual physical duration, affect people's judgments of how long stimuli and events last (Allan, 1979, 1992). For example, moving stimuli have been judged as lasting longer than static ones (Brown, 1995) and familiar words as lasting longer than unfamiliar ones (Witherspoon & Allan, 1985), and a considerable body of work has demonstrated that auditory stimuli usually produce longer subjective durations than do visual ones of the same real length (Goldstone & Lhamon, 1974, summarize studies from the 1950s and 1960s; Wearden, Edwards, Fakhri, & Percival, 1998, and Wearden, Todd, & Jones, 2006, provide more recent discussions). However, probably the best known of these effects is the *filled-duration illusion*.

The literature on the filled-duration illusion shows a pleasing unanimity: Almost all studies found that filled intervals appear to last longer than unfilled ones of the same real-time duration (recent studies that used pigeons [*Columba livia*] by Miki & Santi, 2005, and Santi, Keough, Gagne, & van Rooyen, in press, are exceptions). Apart from these exceptions, however, previous work has been remarkably heterogeneous. One source of disagreement has been the meaning of "filled" and "unfilled." The general reader might suppose that when auditory stimuli were used, for example, filled stimuli probably would be continuous tones, and unfilled ones would be the same durations defined by brief clicks at the beginning and end of the relevant period. Comparisons between such stimulus types have, indeed, been used (e.g., the present study and Rammsayer & Lima, 1991), but some other filled/unfilled duration comparisons use different types of events. In the often-cited article by Thomas and Brown (1974), for example, all intervals were initiated and terminated by click-like stimuli, but filled intervals also contained three other clicks that were regularly or irregularly spaced within the interval. In other work, the filled interval was filled by attending to complex stimuli such as line drawings (Ornstein, 1969).

Another feature of previous work has been the range of experimental and theoretical issues discussed in the context of the filled-duration illusion. If unfilled intervals are started and ended by marker stimuli, for example, does it make a difference how long the markers actually are? Rammsayer and Leutner (1996) showed that it does. When filled intervals are subdivided into periods defined by clicks, how are these subintervals combined? Thomas and Brown (1974) used this method and developed a mathematical treatment of the way in which the combination process produced greater subjective duration estimates in the filled case (see also Adams, 1977). Another issue is that of the influence on time judgments regarding what the stimuli marking the start and end of the unfilled durations actually are, and this has been shown to influence time judgments (see Grondin, Roussel, Lamache, Roy, & Oullet, 2004, for a recent study, and Grondin, 2003, for a recent review). Other experiments have looked at the type of filler material used during the filled interval (e.g., Buffardi, 1971; Foley, Michaluk, & Thomas, 2004). A range of time judgment procedures also has been used in filled/unfilled interval comparisons, and the duration of the intervals used also has varied markedly. At one extreme, Rammsayer and Lima (1991) used a pair-comparison discrimination procedure in which the duration difference between two stimuli presented during a trial varied depending on the accuracy of previous responses, and all durations presented were approximately 50 ms to 60 ms long. In contrast, Thomas and Brown (1974) used a reproduction method and used intervals up to 5,000 ms long, and even longer events have been timed in some other studies (e.g., Ornstein, 1969). Obviously, varying the durations timed by 100-fold or more complicates interpretation of

John H. Wearden, School of Psychology, Keele University; Roger Norton, Simon Martin, and Oliver Montford-Bebb, Department of Psychology, Manchester University.

The experiments reported in this article were conducted by Roger Norton, Simon Martin, and Oliver Montford-Bebb in partial fulfillment of the requirements for bachelor's degrees awarded by Manchester University.

Correspondence concerning this article should be addressed to John H. Wearden, School of Psychology, Dorothy Hodgkin Building, Keele University, Keele, Staffordshire ST5 5BG, England. E-mail: j.h.wearden@psy.keele.ac.uk

results: For example, chronometric counting may intervene at durations of greater than 1 s (and, if it does, it will almost certainly affect timing judgments; see Wearden, Denovan, Fakhri, & Haworth, 1997, for a demonstration), whereas it could not possibly play any role in judgments of 50-ms-long stimuli.

In the present article, we seek to cut the Gordian knot bequeathed to us by previous work. We do this by, first, asking a few apparently straightforward questions about characteristics of what seem to be simple forms of the filled-duration illusion and, second, trying to provide some explanations of the results obtained using ideas from contemporary timing theory. Suppose, in anticipation of results that will be presented later, a filled-duration illusion is obtained; so, for some time value t, judgments of an unfilled duration $U(t)$ are shorter than those of a filled duration $F(t)$. If, in general, $F(t) > U(t)$, how does the effect change as t is varied? For example, is the subjective duration difference between $F(t)$ and $U(t)$ manifested in terms of a constant difference between the two judgments as t is varied (i.e., the difference is additive), or is $U(t)$ always some fraction of $F(t)$ (i.e., the difference is multiplicative)? Such questions have implications for underlying timing mechanisms that might explain the filled-duration illusion, as is shown in the *Discussion* section of Experiment 1.

Another issue of interest is whether filled/unfilled duration perception differences are manifested only in terms of mean judgments or in terms of variability, as well. Filled durations may be perceived as longer than unfilled ones, but are they perceived as relatively more or less variable? One possibility is that unfilled intervals, which have very clearly defined start and end markers, might be perceived as relatively less variable than filled ones, although results in Rammsayer and Lima's (1991) article imply the opposite: In their work, the filled durations produced smaller thresholds, suggesting less variable representations. This issue is addressed in Experiments 2 and 3.

As well as providing data on some simple, but previously somewhat neglected, questions about the filled-duration illusion, another aim of the present article is development of a theoretical treatment of the effects found through use of some ideas derived from a currently popular theory of animal and human timing (Gibbon, Church, and Meck, 1984).

Scalar timing (or *scalar expectancy*) *theory (SET)* has many features in common with an earlier theory of human timing, the *clock model* of Treisman (1963). SET initially was developed as an account of the behavior produced by animals on temporally constrained reinforcement schedules or specially designed timing tasks and continues to enjoy considerable success as an explanation of timing in rats (*Rattus norvegicus*), pigeons, and other animals (e.g., Church, Meck, & Gibbon, 1994; Lejeune, Ferrara, Simons, & Wearden, 1997). For the last 15 years, SET also has been applied to timing in humans. The work of Wearden and McShane (1988) was probably the first study with human participants in which the researchers used SET as an explanatory framework, and since then, a number of studies of human timing through use of the SET framework have appeared (see Allan & Gibbon, 1991, and Wearden, 1991a, 1991b for early examples: Allan, 1998, and Wearden, 2003, provide reviews of work with humans conducted within the SET framework.). This work shows that SET deals well with some aspects of human timing, provided that participants do not use chronometric counting (Wearden, 1991a; Wearden, Denovan, et al., 1997), and the short intervals used exclusively in the present article (ranging from 77 ms to 1,183 ms) are within the range used in previous studies of non-counting-based timing in humans.

For present purposes, the most relevant aspect of SET is its use of a pacemaker–accumulator type internal clock. Here, a pacemaker that produces pulses at some rate (usually assumed as fast) is connected to an accumulator by a switch. When a stimulus that will be timed is presented, the switch closes, allowing pulses to flow into the accumulator, and stimulus offset opens the switch, cutting the flow. The closing and opening operations of the switch need be neither instantaneous nor variance free, as is shown in the *Discussion* section of Experiment 1. Thus, when a stimulus that will be timed ceases, the raw material for judgments about its duration is the number of pulses stored in the accumulator.

In anticipation of results that will be presented later in this article (*Results* section), the stimuli used in this study produced a large filled-duration illusion, with continuous tones being judged as substantially longer than either click-defined intervals or gaps of the same real duration. How can this illusion be explained? The internal clock proposed by SET offers more than one possibility. For example, the pacemaker of the internal clock may run faster for the tones than for the click-defined intervals or the gaps. Alternatively, mean switch latencies or switch operation variances may differ for the two types of events. In the present work, in addition to providing a number of demonstrations of the filled-duration illusion, we try to use the mechanics of the internal clock proposed by SET to explain the effect, thus bringing this venerable temporal "illusion" within the framework of modern timing theory and, in particular, emphasizing its continuity with other phenomena in duration judgment.

Experiment 1

We begin with an initial demonstration of the filled-duration illusion using a *temporal generalization* paradigm. As in all of the experiments reported here, filled intervals were tones produced by the speaker of a conventional computer. In Experiment 1, the unfilled intervals started and ended with click-like stimuli. In the variant of temporal generalization used in some experiments with human participants (e.g., Wearden, 1992; Wearden, Wearden, & Rabbitt, 1997), people initially were presented with a stimulus (e.g., a tone 400 ms long) identified as having a standard duration. Then, they received a series of tones that were shorter or longer than, or equal in duration to, the standard, and they had to decide whether or not each comparison tone had the standard duration. Feedback as to performance accuracy was given after each response.

We modified this method slightly for Experiment 1. There were four types of trial blocks, and as an illustration, we consider the filled/unfilled condition (F/U), in which the first letter indicates the type of stimulus presented as the standard and the second letter indicates the type of stimulus presented as the comparison. The block began with four presentations of a filled interval (a tone), which was identified as having a standard duration. Following this, participants received seven unfilled durations (intervals started and ended with a click-like stimulus). The interval between the clicks was either equal to the standard duration or shorter or longer. After each stimulus was presented, the participant judged whether it had the same duration as the standard, but no feedback was given, as this feedback would have allowed participants to compensate for any filled-duration illusion effect that they demonstrated as the experiment proceeded. The other three conditions, unfilled/unfilled (U/U), filled/filled (F/F), and unfilled/filled (U/F), were identical except for the type of stimuli used as standards and comparisons.

Method

Participants. Sixteen first-year Manchester University undergraduate students participated for course credit, which was not contingent on performance.

Apparatus. Participants were tested individually in a cubicle isolated from external noise and light. An Opus SX-16 IBM-compatible computer with a color monitor controlled all experimental events, and responses were registered on the keyboard. The experimental programs were written in the Micro-Experimental Laboratory (MEL) language (Psychology Software Tools, Inc., Pittsburgh, PA), which assured millisecond accuracy for timing of stimuli and responses.

Procedure. Consider first a block of F/U trials. The block began with four presentations of a 500-Hz tone with a duration that was constant within the block but that was a random value chosen from a uniform distribution running from 400 ms to 600 ms (M = 500 ms), which varied between blocks. Presentations of the standard were separated by gaps chosen at random from a uniform distribution running from 2,000 ms to 3,000 ms. When the standard had been presented, a display informed participants that a number of comparison stimuli would now be presented. Each one was an unfilled interval that started and ended with a 10-ms presentation of a 1,000-Hz tone. We presented seven comparison intervals, the standard, and six nonstandard durations, which we composed by adding the following values to the standard: –300 ms, –200 ms, –100 ms, 100 ms, 200 ms, and 300 ms. That is, intervals from 300 ms below the standard to 300 ms above the standard were presented. The participant delivered each comparison duration by pressing the spacebar in response to a prompt; after a random delay ranging from 2,000 ms to 3,000 ms, this response was followed by the comparison interval. After this task, the participant received the display "Did that stimulus have the same length as the standard? Press *Y* (yes) or *N* (no) keys." No performance-related feedback was given after the response, which was followed by a "Press spacebar for next trial" prompt. The seven comparison durations were arranged in a random order that varied for each participant. The other three blocks were identical except for the relation between the standard and comparison stimuli: In the U/F condition, for example, the standard duration was unfilled, and the comparison durations were filled. In the U/U and F/F blocks, both standard and comparison stimuli were of the same type. The 4 stimulus block types (U/U, F/F, U/F, and F/U) were arranged in a random order that varied for each participant, and the 4 blocks were each presented 4 times, making 16 blocks in all; the order of the conditions within each block of 4 stimulus types was randomly varied between presentations and participants.

Results

The usual method of presenting data from temporal generalization experiments is plotting the proportion of "yes" responses (i.e., identifications of a stimulus as having the standard duration) against stimulus duration. In the present experiment, however, the standard value varied slightly from one block to another; thus, the equivalent method is plotting the proportion of "yes" responses against the comparison–standard difference in ms. Here, 0 indicates the situation in which the standard and comparison stimuli had the same duration, negative values indicate the situations in which the comparison was shorter than the standard, and positive values indicate the situations in which the comparison was longer than the standard. Figure 1 shows the results when plotted in this way.

The upper panel of Figure 1 shows data from conditions in which the standard and comparison stimuli were both filled (F/F) or both unfilled (U/U), and the lower panel shows data from conditions in which the standard and comparison durations were of different types (F/U and U/F). Inspection of the upper panel strongly suggests that when the comparison and standard durations were of the same type, judgments were close to veridical in that the 0-ms difference (i.e., no difference between the standard and comparison duration) produced the highest proportion of "yes" responses. When the standard and comparison stimuli were of different types (lower panel), on the other hand, judgments were systematically distorted, being skewed to the left in the U/F case and to the right in the F/U case.

In an overall analysis of variance (ANOVA) using all four conditions together, we found no significant effect of condition on proportion of "yes" responses, $F(3, 45) = 1.05$, *ns,* but we did find significant effects of stimulus duration, $F(6, 90) = 41.94$, $p < .001$, and a significant Condition × Stimulus Duration interaction, $F(18, 270) = 7.26$, $p < .001$. Of these two significant results, the former merely confirms the fact, obvious from Figure 1, that the proportion of "yes" responses depended on stimulus duration; the latter shows that the different conditions did not produce the same pattern of responding.

In an ANOVA comparing the U/U and F/F conditions, we found no effect of condition overall, $F(1, 15) = 0$, *ns* but a significant effect of stimulus duration, $F(6, 90) = 33.96$, $p < .001$, and a significant Stimulus Duration × Condition interaction, $F(6, 90) = 2.55$, $p = .02$. This latter result suggests that the temporal generalization functions for U/U and F/F have slightly different shapes, as will be examined further in the paragraphs below.

In an ANOVA comparing the U/F and F/U conditions, we found no effect of condition overall, $F(1, 15) = .01$, *ns*, but significant effects of stimulus duration, $F(6, 90) = 9.40$, $p < .001$, and Stimulus Duration × Condition interaction, $F(6, 90) = 7.76$, $p < .001$. The latter result confirms that the response functions from the different conditions have markedly different shapes, as is obvious on inspection of the lower panel of Figure 1.

We used t tests to examine the response functions from the different conditions in more detail. Consider the response functions, shown in the upper panel of Figure 1, from the F/F and U/U conditions. We tested the temporal generalization gradients for asymmetry around the 0-ms difference by comparing the average proportion of "yes" responses from stimulus durations that were shorter than the standard with the average proportion of "yes" responses from stimulus durations that were longer than the standard. When gradients from all four conditions were compared in this way, all were significantly asymmetrical: for F/F, $t = -2.72$; for U/U, $t = -3.61$; for F/U, $t = -3.78$; for U/F, $t = 3.42$; all dfs = 15, all ps $< .05$.

When the gradients from F/U and U/F were compared, a similar analysis showed that they were asymmetrical in different direc-

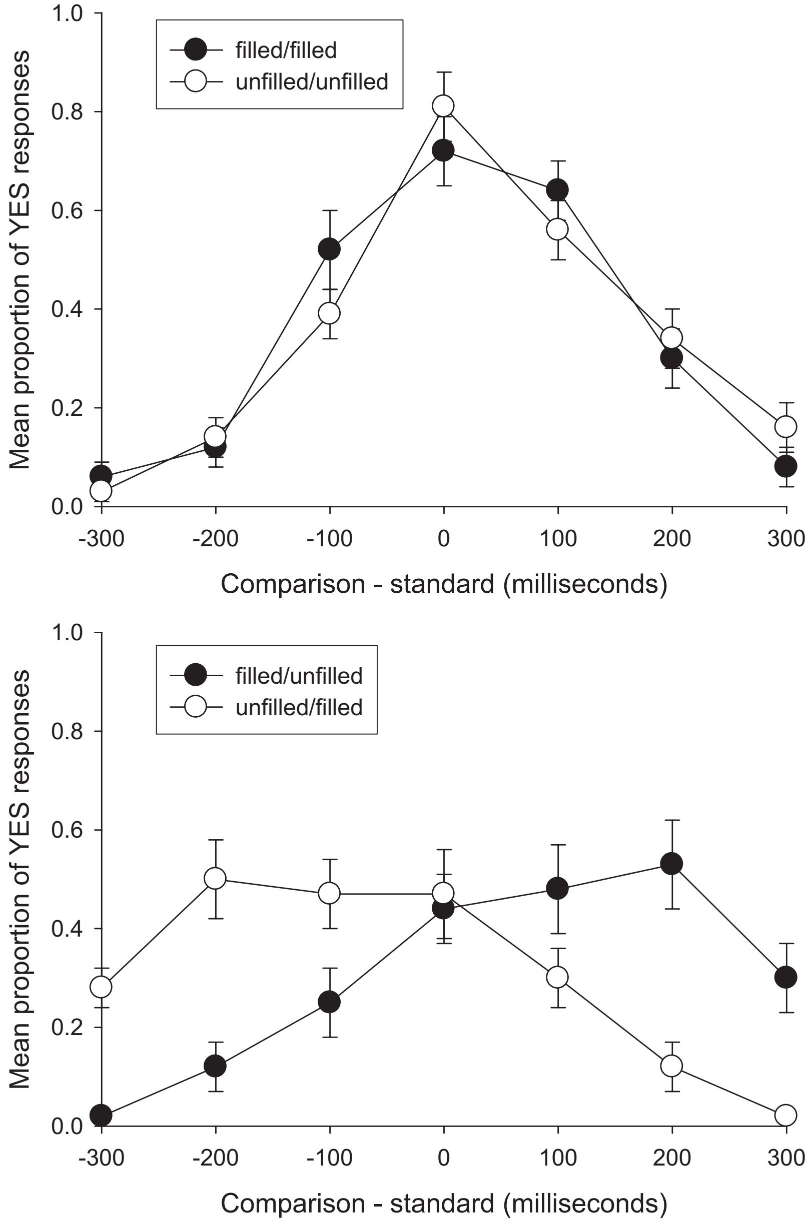

Figure 1. Temporal generalization gradients from Experiment 1. Proportion of "yes" responses (identifications of a comparison stimulus as having the standard duration) plotted against comparison/standard difference. Upper panel: Data from filled/filled and unfilled/unfilled conditions. Lower panel: Data from filled/unfilled and unfilled/filled conditions. Error bars show standard error of the mean.

tions. For example, there were significantly more "yes" responses to comparisons longer than the standard in F/U than in U/F, $t(15) = 3.53$, $p < .01$, and there were more "yes" responses to comparisons shorter than the standard in U/F than in F/U, $t(15) = 3.67$, $p < .01$. This means that when standards were filled and comparisons were unfilled (F/U), comparison stimuli that were longer than the standard were matched to it, whereas in the reverse case (U/F), actually shorter comparison stimuli were maximally identified as the standard. That is, in the F/U and U/F comparisons, the filled intervals appeared significantly longer than the unfilled ones, a demonstration of the filled-duration illusion.

Discussion

Experiment 1 demonstrated the filled-duration illusion with the stimuli used in our study, and it should be noted that the effect was obtained even though if the participants had used the whole of the unfilled stimulus (i.e., starting from the onset of the first click to the offset of the second rather than just timing the interval between, as instructed), the unfilled stimulus actually would have been 20 ms longer than the filled one. Results in the lower panel of Figure 1 also show that the filled-duration illusion was very marked; for example, in the F/U condition, a comparison duration

200 ms longer than the standard duration was maximally matched to it, and in the U/F condition, a comparison duration 200 ms shorter than the standard duration was maximally chosen. Given that the average standard duration was 500 ms, the illusion was large in percentage terms, apparently larger than the auditory/visual difference investigated by Wearden et al. (1998) through the use of temporal generalization (cf. Wearden et al., 1998, Figure 1, p. 102, and the General Discussion later in this article).

The filled-duration illusion depended, however, on participants experiencing both types of stimulus in the same trial block; when stimulus types were the same (U/U and F/F, upper panel of Figure 1), the peak of "yes" responses was at the 0-ms difference. The U/U and F/F temporal generalization gradients were both slightly asymmetrical (with more "yes" responses to comparisons that are longer than the standard than to comparisons that are shorter than the standard), but this asymmetry is normal for temporal generalization in humans (Wearden, 1991a, 1992; Wearden, Denovan, et al., 1997; Wearden, Wearden, & Rabbitt, 1997; for a possible theoretical explanation, see Wearden, 2004), so data from U/U and F/F are consistent with those from other studies even when these previous experiments used slightly different procedures.

Experiment 1 establishes that a strong filled-duration illusion is present with the stimuli used but tells us little more about the effect. Why are the filled stimuli perceived as longer than the unfilled ones? According to quantitative accounts of the operation of internal clocks, subjective duration differences between stimuli, such as those obtained in Experiment 1, can arise in a number of ways. Consider a simple pacemaker–accumulator internal clock (e.g., Gibbon & Church, 1984), and suppose that the operation of this clock differs for filled and unfilled intervals. One possibility is that pacemaker rate varies for the different stimuli, running at rate r_f for filled stimuli and r_u for unfilled ones, with $r_f > r_u$. Therefore, for some constant interval length *t,* more pulses will accumulate for filled than unfilled intervals, giving rise to a subjective duration difference in the direction found in Experiment 1. But this is not the only possibility; another concerns the latency of operation of a hypothesized switch between the pacemaker and accumulator. According to Gibbon and Church (1984), when a stimulus that will be timed begins, pulses flow from the pacemaker to the accumulator through a switched connection, but this switch has nonzero latency l_c for closing (i.e., allowing pulses to flow) and l_o for opening (and cutting the pacemaker–accumulator connection) when the stimulus ceases. Thus, for a pacemaker rate *r*, the number of pulses accumulating in some time *t* is $r(t - l_c + l_o)$. Subjective differences in duration between filled and unfilled intervals could occur if the balance of switch latencies differed between the stimulus types (e.g., the switch closed faster or opened more slowly with filled rather than unfilled stimuli), even if pacemaker rate *r* was constant.

Inspection of the function for pulse accumulation suggests, however, that pacemaker speed effects and switch effects might, in some cases, be dissociated. The function can be divided into two additive components, $rt + r(l_o - l_c)$, the first of which varies as the duration timed varies and the second of which is a multiple of pacemaker rate and the difference between the latencies of opening and closing the switch of the accumulator but does not depend on the duration timed *t*. If pacemaker rate *r* varies between filled and unfilled intervals, differences would be expected both in the slope of the function relating estimates to *t* (the first component) and in the intercept (the second component), but observing the latter effect would depend on the difference $l_o - l_c$ being greater than zero. Even if the absolute values of switch opening and closing were different for filled and unfilled intervals (so, e.g., both the onset and offset of stimuli that would be timed were registered more rapidly for filled than for unfilled stimuli), an intercept difference between the stimulus types would be observed only if the difference between opening and closing latencies varied between the stimulus types. In contrast, a constant difference in perceived duration between filled and unfilled intervals, independent of the interval timed, would suggest that the difference in subjective duration estimates arises because of switch latency effects; for example, the switch closes more rapidly, or opens more slowly, in the filled than in the unfilled case, so a constant number of extra pulses accumulates during the former stimulus type.

Experiment 2

In Experiment 2, we used verbal estimation of 10 different durations (ranging from 77 ms to 1,181 ms). It consisted of two subexperiments that were, procedurally, almost identical except for the type of stimulus used in the respective unfilled condition; therefore, we report these two subexperiments together to save space. In both subexperiments, 500-Hz tones served as filled stimuli (as in Experiment 1), and in Experiment 2A, the unfilled stimuli were defined by click-like markers—again, as in Experiment 1. For simplicity, we refer to this as the *click version* of the F/U effect. In Experiment 2B, in contrast, we used gaps in 500-Hz tones as the unfilled stimuli. On these trials, a tone was presented for a short random duration, then a gap ensued, then the tone recommenced for another short, random time. The participant's task was estimation of the duration of the gap. We refer to this procedure as the *gap version* of the filled-duration effect. As noted earlier, stimuli used in demonstrations of the filled-duration illusion have varied markedly from one study to another, and our intention in comparing two different sorts of unfilled stimuli (click-defined and gaps) was simply investigating whether similar F/U differences would occur with both stimulus types.

In both Experiment 2A and Experiment 2B, the participants experienced intermixed presentations of both the filled and the unfilled intervals and provided verbal estimates of their duration in ms. A similar range of stimuli was used in three other studies, one on speeding up the internal clock (Penton-Voak, Edwards, Percival, & Wearden, 1996) and the other on modality differences in duration judgments (Wearden et al., 1998, 2006). In both studies, researchers showed that the stimulus range used was sufficient for deciding, by means of regression techniques, whether perceived duration differences were multiplicatively related to duration (i.e., pacemaker speed) or constant (i.e., a switch effect); thus, the same range was used here.

Method

Participants. Fifteen (Experiment 2A: click version) and 17 (Experiment 2B: gap version) undergraduates at Manchester University participated for course credit.

Apparatus. Apparatus was the same as that used in Experiment 1.

Procedure. Participants experienced a single experimental session. The stimulus durations that would be estimated were 77 ms, 203 ms, 348 ms, 461 ms, 582 ms, 707 ms, 834 ms, 958 ms, 1065 ms, and 1181 ms. In both Experiment 2A and Experiment 2B, the filled intervals were continuous 500-Hz tones of one of the durations given above, produced by the computer speaker. In Experiment 2A, the unfilled intervals commenced with a 1,000-Hz, 10-ms tone produced by the computer speaker, then a delay of one of the values above, then a second 1,000-Hz 10-ms tone. In Experiment 2B, the unfilled intervals were defined as follows: A 500-Hz tone started and ran for a random value selected from a uniform distribution running between 300 ms and 500 ms, then a gap lasting one of the values given above ensued, then a second tone lasting for a random value between 300 ms and 500 ms ensued. In both subexperiments, the 20 stimuli (filled and unfilled intervals of the 10 durations given above) were arranged in a random order and were presented once each in a block. A different random order was then selected for subsequent blocks, until five blocks (100 trials) had been presented. In Experiment 2B, a visual display (the words "tone" or "gap") preceded each stimulus to inform participants which duration should be estimated, but in Experiment 2A, the stimuli (which were obviously distinguishable) were intermixed without any display.

We provided participants with appropriate instructions about which durations should be estimated (tones, gaps between clicks, or gaps in tones), and we required them to type their verbal estimate of stimulus duration in ms using a scale in which 1,000 = 1 s. They were informed that all durations were between 50 ms and 1,500 ms. In both subexperiments, the participant delivered the stimulus that was going to be estimated by pressing the spacebar, and the stimulus started between 2,000 ms and 3,000 ms after the response. No feedback as to estimate accuracy was given.

Results

Experiment 2A: Click version. The upper panel of Figure 2 shows mean verbal estimates plotted against stimulus duration for the filled and unfilled click conditions, the center panel shows standard deviations of estimates, and the lower panel shows coefficients of variation (standard deviation/mean) from the same data. We filtered out verbal estimates of less than 50 ms and greater than 1,500 ms to correct the data for typographical errors, but this filtering rejected only a few observations.

Inspection of the upper panel of Figure 2 suggests that the filled intervals produced longer verbal estimates than the unfilled ones and that the difference between the two was greater at longer durations than at shorter ones, that is, was multiplicative with duration. An ANOVA of means found significant effects of condition (filled vs. unfilled), $F(1, 14) = 57.96$, $p < .001$; stimulus duration, $F(9, 126) = 128.56$, $p < .001$; and Stimulus Duration × Condition interaction, $F(9, 126) = 10.01$, $p < .001$. The first of these effects shows an overall filled-duration illusion, the second confirms the result obvious from inspection of Figure 2 that mean estimates varied with stimulus duration, and the third suggests that the F/U difference was a slope effect, that is, was consistent with the idea that the pacemaker of an internal clock ran at different rates for the filled and unfilled conditions.

Whether the F/U perceived duration difference was a slope or intercept effect was also tested in another way. Verbal estimates from individual participants in the filled and unfilled conditions were regressed against stimulus duration, and a slope and intercept value for each individual was calculated. Values from the different stimulus conditions were then compared by *t* tests. Mean slope

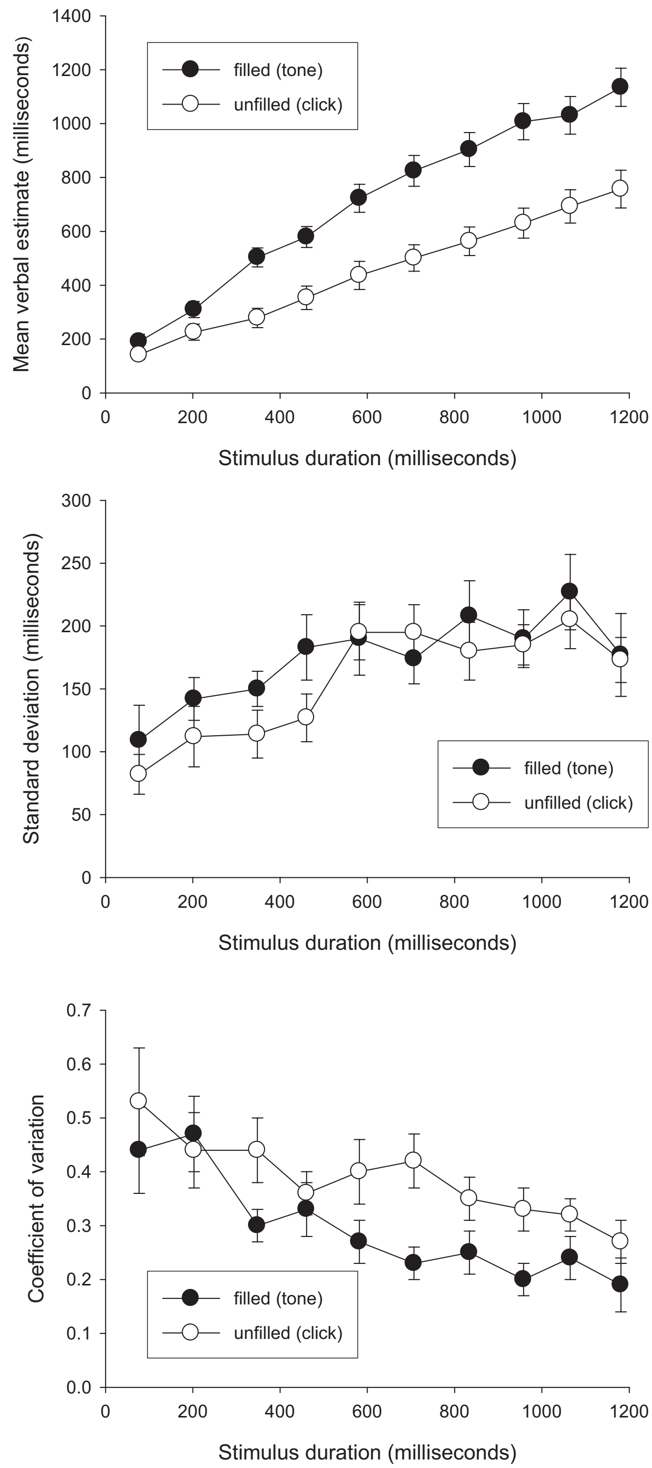

Figure 2. Mean (upper panel), standard deviation (center panel), and coefficient of variation (standard deviation/mean; lower panel) of verbal estimates plotted against stimulus duration from the click version of the filled-duration illusion (Experiment 2A). Error bars show standard error of the mean.

from the filled condition was 0.85 and from the unfilled condition was 0.55, and this difference was significant, $t(14) = 4.41$, $p < .01$. In fact, 14 of 15 participants had higher slopes for the filled variables than for the unfilled values. Intercepts in ms were 176.2

(filled) and 101.5 (unfilled), and this difference was also significant, $t(14) = 2.68$, $p < .05$.

Inspection of the center panel of Figure 2 suggests that although standard deviations of estimates increased with increasing stimulus duration, there were no marked differences between the values produced from filled and unfilled intervals. These suggestions were confirmed by an ANOVA, with no significant effect of condition (filled vs. unfilled), $F(1, 14) = 1.96$, *ns,* and a significant effect of stimulus duration, $F(9, 126) = 5.20$, $p < .001$, but no significant Condition × Stimulus Duration interaction, $F(9, 126) = 0.64$, *ns*. However, although estimates from the filled and unfilled conditions did not differ in absolute variability, they did differ in relative variability, as the lower panel of Figure 2 shows. Here, the coefficient of variation statistic (standard deviation/mean) expresses relative variability of estimates, independent of any mean differences that may be present. Inspection of the data suggests that unfilled intervals produced relatively more variable estimates than do filled ones, and another effect that seems to be present is a decline in coefficient of variation as stimulus durations increased (i.e., coefficients of variation were smaller at longer durations than at shorter ones). Consistent with this finding, an ANOVA found a significant F/U difference in coefficient of variation overall, $F(1, 14) = 15.24$, $p < .001$, indicating that unfilled intervals produced significantly more variable estimates than did filled ones, and a significant effect of stimulus duration, $F(9, 126) = 5.18$, $p < .001$, showing that coefficient of variation declined significantly with increasing stimulus duration. The Stimulus Duration × Condition interaction was, however, not significant, $F(9, 126) = 1.22$, *ns,* suggesting that the decline in coefficient of variation with increasing stimulus duration was similar for judgments of both filled and unfilled intervals.

Experiment 2B: Gap version. The upper panel of Figure 3 shows mean verbal estimates resulting from the tones and gaps. Inspection of the data suggests that the tones produced longer verbal estimates than did the gaps, with the difference between the two increasing with increasing stimulus duration. ANOVAs found significant effects of stimulus type, $F(1, 16) = 64.89$, $p < .001$; stimulus duration, $F(9, 144) = 52.26$, $p < .001$; and Stimulus Type × Duration interaction, $F(9, 144) = 18.97$, $p < .001$. The first of these findings shows that the filled-duration illusion was manifested in the gap version of our task, the second finding shows that mean estimates grew with stimulus duration, and the third finding shows that the filled duration effect was multiplicative with duration, that is, consistent with a pacemaker speed, rather than switch latency, interpretation.

As for the click version of the task, linear regression of verbal estimates against stimulus duration was conducted for individual participants, and slopes and intercepts from the tone and gap conditions were compared. Mean slope from the tone condition was 1.0, and mean slope from the gap condition was 0.52; these differed significantly, $t(16) = 9.60$, $p < .001$. In fact, all 17 participants showed a higher regression slope value in the tone condition than in the gap condition. Mean intercepts were 13.65 ms for the tones and 65.65 ms for the gaps, and these values just reached a significant difference, $t(16) = -2.23$, $p < .05$.

The center panel of Figure 3 shows standard deviations of estimates, and inspection suggests that values from the filled and unfilled intervals showed little obvious difference but that both increased with increasing stimulus duration. This finding was confirmed by an ANOVA, which found that the only significant effect was stimulus duration, $F(9, 144) = 16.94$, $p < .001$. Neither the effect of condition (filled vs. unfilled) nor the Condition × Stimulus Duration interaction approached significance: condition, $F(1, 16) = 0.06$, *ns*; interaction, $F(9, 144) = 1.47$, *ns*.

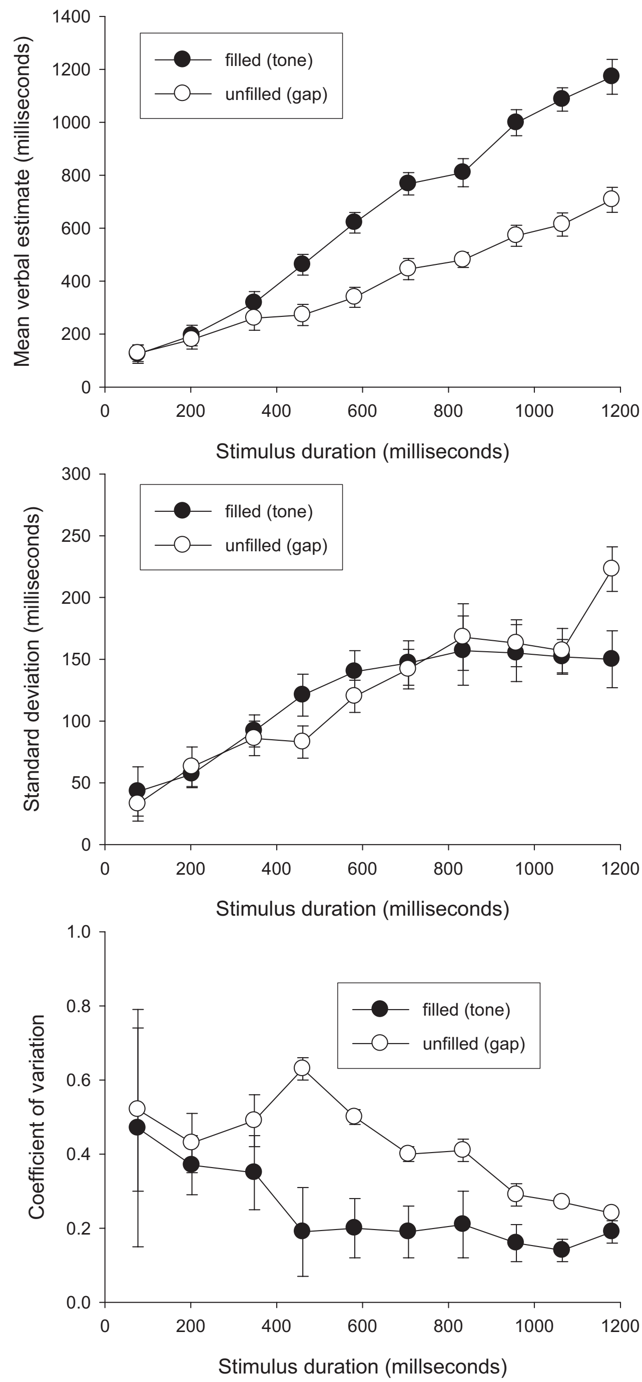

Figure 3. Mean (upper panel), standard deviation (center panel), and coefficient of variation (standard deviation/mean; lower panel) of verbal estimates plotted against stimulus duration from the gap version of the filled-duration illusion (Experiment 2B). Error bars show standard error of the mean.

Upon inspecting variation coefficients of estimates from the gap version of the task (lower panel of Figure 3), we found that the gaps produced relatively more variable estimates than did the tones and that coefficient of variation value declined with increasing stimulus duration. Both suggestions were confirmed statistically: gaps, $F(1, 16) = 7.91$, $p = .01$; tones, $F(9, 144) = 2.11$, $p = .03$; however, the Stimulus Type × Duration interaction was not significant, $F(9, 144) = .64$, *ns*, indicating that the decline in coefficient of variation with increasing stimulus duration was similar for both tones and gaps.

Discussion

The results from Experiment 2 can be simply summarized. Filled intervals produced significantly longer verbal estimates of duration than did unfilled ones, whether the intervals were defined by clicks (Experiment 2A) or gaps in tones (Experiment 2B). In both subexperiments, the F/U mean estimate difference increased with increasing stimulus duration, as evidenced by significant ANOVA Estimate × Duration interactions and by the significantly higher slopes derived from regression analyses of the filled intervals.

Standard deviations of estimates from filled and unfilled conditions did not differ significantly in either Experiment 2A or Experiment 2B, although unfilled intervals produced verbal estimates that were relatively more variable in both cases and coefficients of variation that declined as stimulus duration increased. Given the absence of any difference between filled and unfilled intervals in standard deviations, the obvious interpretation of the coefficient of variation difference is that it is due to the large difference in mean estimates.

The decline in coefficient of variation with increasing stimulus duration is characteristic of data obtained with a verbal estimation method (see Wearden, 1999, and 2003, for examples) and constitutes a violation of the s*calar property* of time, essentially the requirement that the coefficient of variation remain constant with interval value (see Lejeune & Wearden, 2006, for more precise definition and discussion of both properties). Why data from verbal estimation routinely violate the scalar property when the property holds for identical stimuli (e.g., short-duration tones) used in other procedures (e.g., temporal generalization; Wearden, 1992) remains a matter for conjecture. Wearden (2006) presented a computer model of the verbal estimation of durations in the same range as those used here that showed that, at least in some cases, declining coefficients of variation with increasing stimulus duration with this procedure could be attributed to participants' very marked tendency to use only certain estimate values (e.g., those ending in "00" or, more rarely, "50"). In any case, the decline of coefficient of variation with increasing stimulus duration for the click and gap stimuli used in Experiment 2 emphasizes their similarity with filled auditory and visual stimuli, with which this effect is also found.

In Experiment 2, we aimed to distinguish between interpretations of the filled-duration illusion on the basis of pacemaker speed and those due to switch effects. The former would produce slope differences and (possibly) small intercept differences, with the filled condition having larger values of both slope and intercept, whereas the latter interpretation would produce an intercept but no slope effect, again favoring the filled duration. Further consideration of the application of clock model mechanics is provided in the General Discussion, but for present purposes, we note only that significant slope effects were obtained with both the click and gap form of the filled-duration illusion. The click form, in fact, produced exactly the pattern of results predicted, that is, higher slope and intercept with filled durations. The gap form, on the other hand, although demonstrating a slope effect, in addition showed an intercept effect in the opposite direction (i.e., greater intercept for gaps than for tones).

In general, then, mean verbal estimate differences between filled and unfilled intervals seem consistent with a faster pacemaker speed occurring in the filled condition. Does this mean that there is a single pacemaker governing processing of both stimulus types or that there are two pacemakers with intrinsically different speeds? If two pacemakers are operating, they appear to have some similar properties in that both generate mean verbal estimates that increase in an approximately linear manner with stimulus duration. In Experiment 3, we further explored the apparent similarity between unfilled and filled intervals by examining potential effects of speeding up the pacemaker of the internal clock. Penton-Voak et al. (1996) showed that preceding tones and visual stimuli with a train of clicks increased judgments of stimulus duration (see also Treisman, Faulkner, Naish, & Brogan, 1990), and in Experiment 3, we investigated this issue with filled and unfilled intervals.

Experiment 3

In Experiment 3, Penton-Voak et al.'s (1996) method of preceding stimuli that would be estimated with a train of clicks was applied to filled and unfilled durations (tones as well as the same intervals defined by clicks). Three previous studies (Penton-Voak et al., 1996; Wearden et al., 1998; Wearden, Philpott, & Win, 1999) show that this manipulation increased the subjective duration of auditory stimuli, so an effect on filled intervals was expected. It was, however, unclear whether unfilled intervals could be manipulated in this way. Penton-Voak et al. (1996) showed that an unfilled interval produced by a participant showed "speeding-up-the-clock" effects, but estimation of unfilled intervals has not been used previously with this manipulation.

In Experiment 3, we asked two basic questions. First, would both filled and unfilled intervals be affected similarly (or at all) by Penton-Voak et al.'s (1996) "speeding-up-the-clock" manipulation? If they were both affected similarly, this would obviously support the view that filled and unfilled interval estimates were being generated either by the same pacemaker or two very similar ones. Second, if mean verbal estimates could be manipulated, how would variability of estimates change, if at all?

In Experiment 3, participants received filled and unfilled intervals, which had six different durations, and these two stimulus types were sometimes preceded by a 5-s train of clicks and, sometimes, by silence. We distinguished the clicks in the click train from those that we used to define the unfilled interval by ensuring that the former were of different pitch and duration.

Method

Participants. Nineteen Manchester University undergraduates participated for course credit.

Apparatus. Apparatus was the same as that used in Experiments 1 and 2.

Procedure. All participants received a single experimental session. The filled and unfilled intervals had six different durations: 77 ms, 348 ms, 582 ms, 767 ms, 958 ms, and 1,183 ms.

Filled intervals were 2,000-Hz tones produced by the computer speaker. Unfilled interval started with a 15-ms, 2,000-Hz tone, then was followed by a gap chosen from one of the six duration values, and then was followed by another 15-ms, 2,000-Hz tone. On half of all trials, the stimulus (filled or unfilled) was preceded by a 5-s train of 5-Hz clicks; on the other half, the stimulus was preceded by 5 s of silence. The clicks in the click train had a different frequency (500 Hz) and different length (10 ms) from those of the clicks that we used to define the unfilled intervals and were perceptually different from them. They were presented every 200 ms, onset to onset. A typical trial (clicks, filled interval) proceeded as follows. In response to a "Press spacebar for next trial" prompt, the participant pressed the spacebar, and this was followed by 5 s of clicks, as previously defined. The click train was immediately followed by a filled interval that lasted for one of the six durations given previously in this section and then was followed by a prompt requiring verbal estimation, conducted as in Experiment 2. Participants were told that all stimulus durations were between 50 ms and 1,500 ms. Trials of other types (no-clicks and unfilled) were similar; in the no-click conditions, the stimulus was preceded by 5 s of silence, and the unfilled intervals were defined as outlined above. The combination of six stimulus durations, filled and unfilled intervals, and clicks and no clicks generated a basic block of 24 trials. The trials within the block were presented to the participant in a random order (so that any trial type could follow any other) that differed between participants and blocks, and four blocks (96 trials) were given in the session.

Results

As for Experiment 2, we filtered the data to discard the few estimates outside the range of 50 ms to 1,500 ms specified to participants. The upper panel of Figure 4 shows mean verbal estimates from the four different conditions (filled/clicks, filled/no clicks, unfilled/clicks, unfilled/no clicks). Inspection of the data suggests that (a) there was a large F/U difference, with higher verbal estimates in the filled case, both with and without clicks, and (b) the clicks produced a small but consistent increase in the mean verbal estimates of both filled and unfilled stimuli.

ANOVAs of the mean verbal estimates produced a large number of significant results. To simplify exposition, we note that all ANOVAs of mean verbal estimates produced a significant effect of stimulus duration (smallest $F > 100$, $p < .001$). This merely shows that verbal estimates were sensitive to stimulus duration and will not be mentioned further. We concentrate, instead, on effects of stimulus condition (F/U), clicks (present or absent), and interactions between these variables and stimulus duration.

An overall ANOVA found significant effects of the presence or absence of clicks, $F(1, 18) = 26.64$, $p < .001$; a significant effect of stimulus type (filled vs. unfilled), $F(1, 18) = 43.74$, $p < .001$; and a significant Stimulus Type × Stimulus Duration interaction, $F(5, 90) = 19.76$, $p < .001$; however, no significant effects were found for the two-way Clicks × Stimulus Type interactions, $F(1, 18) = 0.17$, *ns*; the two-way Clicks × Stimulus Duration interactions, $F(5, 90) = 1.43$, *ns*; or the three-way Clicks × Stimulus Type × Duration interaction, $F(5, 90) = 1.33$, *ns*. Although these results show both effects of clicks and a filled-duration illusion, the overall data are probably more easily understood with some simpler comparisons.

Was an F/U perceived duration difference found both with and without clicks? For stimuli without clicks, an ANOVA found a significant stimulus type difference, $F(1, 18) = 41.0$, $p < .001$, and a significant Stimulus Type × Duration interaction, $F(5, 90) = 12.81$, $p < .001$. The same result was obtained with clicks: stimulus type, $F(1, 18) = 43.47$, $p < .001$; interaction, $F(5, 90) = 14.34$, $p < .001$.

Did clicks have an effect? For filled intervals, clicks significantly increased mean estimates, $F(1, 18) = 12.44$, $p < .01$, but the Clicks × Duration interaction was not significant, $F(5, 90) = 1.20$, *ns*. The same pattern was obtained with unfilled intervals: click effect, $F(1, 18) = 14.50$, $p < .01$; interaction, $F(5, 90) = 1.69$. In addition, the overall Nonsignificant Clicks × Stimulus Type interaction, reported above, showed that the effect of clicks was the same on both stimulus types.

As in Experiment 2, we regressed verbal estimate against stimulus duration for individual participants for all four conditions, and we performed some comparisons (using t tests) on the resulting slopes and intercepts. Mean slopes were as follows: filled/no clicks, .89; filled/clicks, .96; unfilled/no clicks, .51; unfilled/clicks, .59. Corresponding intercept values were 155.05 ms, 162.84 ms, 73.21 ms, and 68.53 ms. Comparison of filled and unfilled intervals showed that 17 of the 19 participants showed higher slope values in the filled case both with and without clicks, and both differences were significant: no clicks, $t(18) = 5.63$; clicks, $t(18) = 5.09$, ps $< .001$. Fourteen of the 19 participants showed larger slopes after clicks when intervals were unfilled, $t(18) = 2.50$, $p < .05$, whereas 11 of the 19 participants showed higher slopes after clicks when intervals were filled, a difference that did not reach significance, $t(18) = 1.41$, *ns*. Clicks had no significant effects on intercepts, either for filled or unfilled intervals, but filled intervals produced significantly higher intercepts than did unfilled ones with clicks, $t(18) = 3.45$, $p < .01$; however, the difference with clicks did not reach significance, $t(18) = 1.99$, $p = .06$.

In summary, therefore, data from Experiment 3 suggested that the mean estimate difference from the filled and unfilled intervals, observed both with and without clicks, was consistent with a pacemaker speed difference (i.e., higher slope) in the filled condition. Alternatively, the effect of clicks was more ambiguous, producing a significant slope effect for unfilled intervals but a nonsignificant one for filled intervals.

The center panel of Figure 4 shows standard deviations of estimates from the four conditions. Inspection suggests no clear overall differences in standard deviations among the four conditions but an increase in all cases with increasing stimulus duration and the possibility that, at shorter durations, the filled conditions produce more variable estimates. These suggestions were confirmed by an ANOVA that found no overall effect of condition (filled or unfilled, clicks or no clicks), $F(3, 54) = 1.94$, *ns*, but a significant effect of stimulus duration, $F(5, 90) = 15.21$, $p < .001$, and a significant Condition × Stimulus Duration interaction, $F(15, 270) = 1.92$, $p = .02$. Simpler ANOVAs suggested that the source of the interaction was F/U comparisons. When comparing filled and unfilled intervals without clicks, we found that the interaction approached significance, $F(5, 90) = 2.05$, $p = .08$, and, with clicks, was just significant, $F(5, 90) = 2.32$, $p < .05$, whereas comparisons of filled and unfilled intervals with and without clicks yielded F values that were far from significant, $F(5, 90) = .45$ (filled), *ns*, and $F(5, 90) = .90$ (unfilled), *ns*.

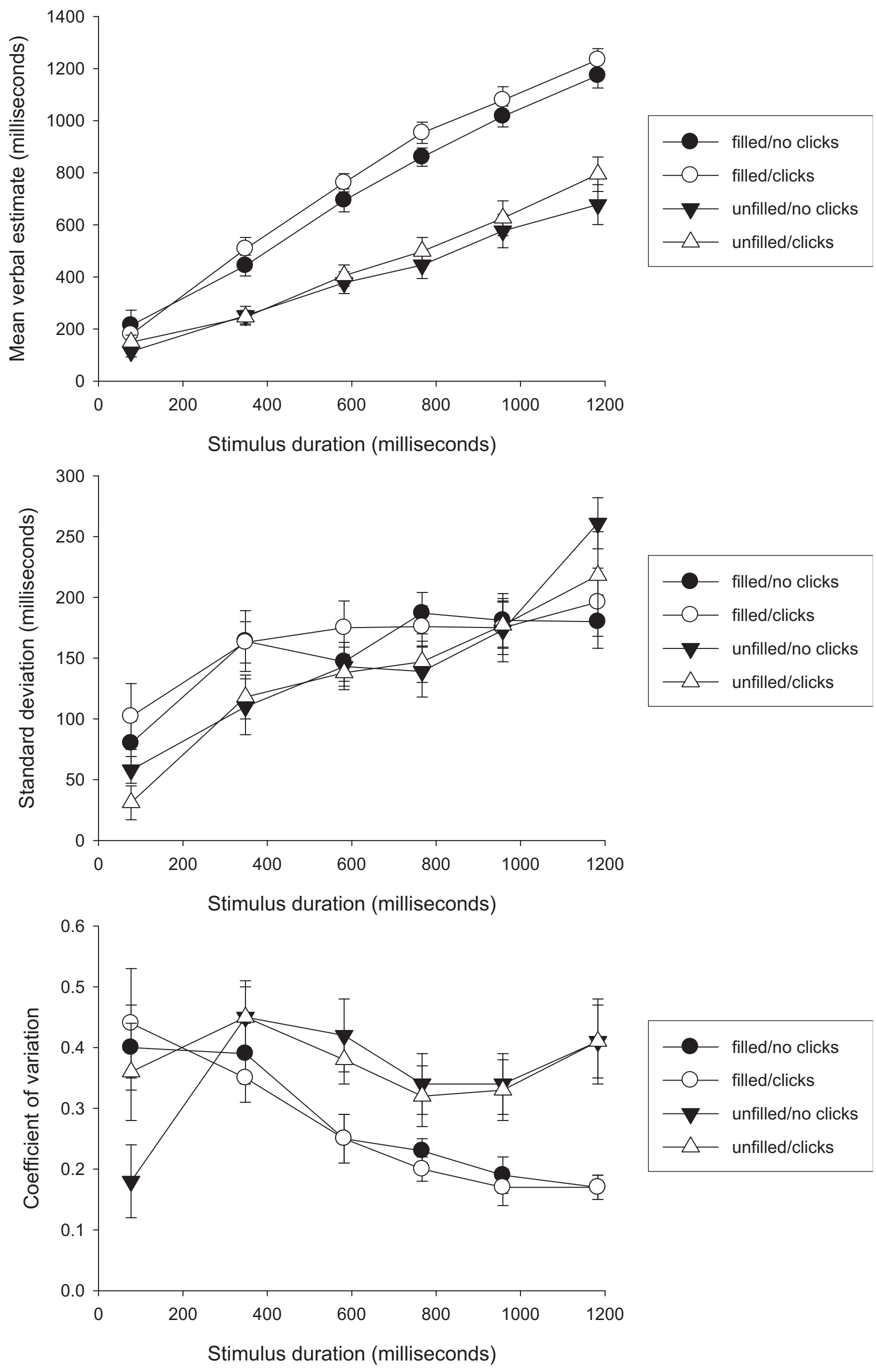

Figure 4. Mean (upper panel), standard deviation (center panel), and coefficient of variation (lower panel) of verbal estimates plotted against stimulus duration for the four conditions of Experiment 3. Error bars show standard error of the mean.

The lower panel of Figure 4 shows coefficients of variation of verbal estimates from the four conditions. Inspection suggests that (a) filled intervals produced smaller coefficients of variation than did unfilled ones; (b) clicks had no very consistent effect, although they produced some reduction in coefficient of variation compared with the identical stimulus type without clicks; and (c) coefficients of variation declined with increasing stimulus duration, at least in the filled condition.

An overall ANOVA of coefficients of variation found significant effects of stimulus type, filled or unfilled, $F(1, 18) = 7.12$, $p < .02$, and stimulus duration, $F(5, 95) = 3.85$, $p < .01$, but no effect of clicks, $F(1, 18) = .09$, *ns*. The presence (or absence) of Clicks × Stimulus Duration interaction just reached significance, $F(5, 90) = 2.49$, $p = .04$, and the Stimulus Type × Duration interaction was significant, $F(5, 90) = 6.25$, $p < .001$, but the Clicks × Stimulus Type interaction was not, $F(1, 18) = .45$, *ns,* and neither was the three-way Stimulus Type × Clicks × Duration interaction, $F(5, 90) = .63$, *ns*. However, inspection of the lower panel of Figure 4 suggests that data from shortest stimulus duration used (77 ms) were somewhat anomalous, with an abnormally low coefficient of variation value from the unfilled/no clicks condition, for example. When the overall ANOVA was rerun with the 77-ms stimulus eliminated, all previously significant interactions disappeared, leaving as the only significant effects the F/U difference and stimulus duration.

Some simpler ANOVAs confirmed this general picture of a lack of effect of clicks on coefficient of variation and a large effect of the F/U difference. For example, clicks had no significant effect on coefficient of variation with either filled intervals, $F(1, 18) = .12$, *ns,* or unfilled intervals, $F(1, 18) = 0.52$, *ns,* whereas filled intervals produced smaller coefficients of variation either without clicks, $F(1, 18) = 5.69$, $p < .03$, or with them, $F(1, 18) = 5.97$, $p < .03$.

Discussion

Data from Experiment 3 confirmed and extended the conclusions drawn from the results of Experiment 2. Filled intervals, whether preceded by clicks or not, produced longer verbal estimates than did unfilled ones, and the analysis of regression slopes was consistent with the F/U difference being attributable to faster pacemaker speed in the filled case. There were no significant differences between filled and unfilled intervals in terms of absolute variability (standard deviation), but relative variability (coefficient of variation) was smaller for filled intervals. Clicks had no effect on either absolute or relative variability. As in Experiment 2, it seems that the cause of differences in coefficients of variation between judgments of filled and unfilled intervals was differences in mean estimates.

General Discussion

Our three experiments taken together confirmed, in various ways, the venerable finding that filled intervals are perceived as, on average, longer in duration than are unfilled ones. In contrast, we found no differences in absolute variability between verbal estimates of filled and unfilled intervals. This finding suggests that there is a degree of independence between the processes responsible for mean judgments and those that determine variability, and one of these factors can differ markedly between conditions, whereas the other factor can show little or no change. The question of what determines variance of timed behavior in different tasks is a vexing one (see Gibbon & Church, 1984; Jones & Wearden, 2003, 2004, and Wearden & Bray, 2001, for discussions), and no published model of determinants of variance on verbal estimation tasks currently exists.

In addition, the fact that standard deviations obtained from estimates of filled and unfilled intervals never differed when the same time values were estimated, whereas coefficients of variation always did, raises the question of how variability, in this sort of case, should be interpreted. Performance on many timing tasks conforms to the two scalar properties of timing (see Lejeune & Wearden, 2006, for a review). One of these scalar properties is *mean accuracy*, the requirement that mean measures of timed behavior vary accurately with the time requirement of the task. The second is the scalar property of *variance*, the requirement that the coefficient of variation of measures of timed behavior remains constant as the interval timed varies.

If the mean accuracy property holds (and Lejeune & Wearden, 2006, show that it usually does), then if two different conditions differ in standard deviation (absolute variability) at some particular time value, they also will differ in coefficient of variation (relative variability), as the means will be the same or nearly so. However, in the case of verbal estimation, mean judgments are not accurate, and mean estimates from different conditions with the same real-time durations, such as filled and unfilled intervals, may vary markedly; therefore, plotting standard deviation against real stimulus duration (as in the center panels of Figures 2, 3, and 4), as well as constructing a coefficient of variation and plotting it against stimulus durations (as in the lower panels of Figures 2, 3, and 4), provide different views of variability.

The presence of the mean accuracy property has made relative measures of variability (coefficients of variation, and Weber fraction–like measures) nearly ubiquitous in studies of timing related to the SET framework, and the coefficient of variation is often used as an index of timing sensitivity (e.g., Lejeune & Wearden, 1991). Our data show that, in this sense, the timing of unfilled intervals is less sensitive than is the timing of filled ones. The F/U difference found here resembles that difference previously found when auditory and visual stimuli were used in duration comparisons. In this case, auditory stimuli are perceived as longer and have smaller coefficients of variation, compared with visual stimuli (Wearden et al., 1998). However, the mean duration difference between filled and unfilled intervals found in these experiments is considerably greater in magnitude than the auditory/visual difference found in a number of previous experiments. For example, if we consider the data from Experiments 2 and 3, the average estimate for a click-defined stimulus in Experiment 2 was 63.5% of the estimate for a tone, and the comparable value for the gap stimulus was 60.9%. In Experiment 3, the U/F values were 55.7% and 57.6% for stimuli with and without clicks, respectively. In contrast, it has been found that visual stimuli have been subjectively judged as approximately 80%–90% of the duration of auditory stimuli, and similar results are found in verbal estimates in Wearden et al. (1998, 2006) and from values used in theoretical models in Droit-Volet, Tourret, and Wearden (2004) and Penney, Gibbon, and Meck (2004). Contrasting with both of these is the effect of clicks on mean estimates, which, although statistically significant, is very small in percentage terms (in Experiment 3 of this study, the percentages were 7% and 11%, respectively, for filled and unfilled intervals, and similar values were obtained in Penton-Voak et al., 1996).

The results from mean verbal estimates in the present study comparing different sorts of filled and unfilled intervals and from comparisons of judgments of the duration of auditory and visual stimuli in Wearden et al. (1998, 2006), Penney et al. (2000), and Droit-Volet et al. (2004, in a study demonstrating auditory/visual duration judgments in children) are all compatible with differences in pacemaker speed, with rates being higher in filled and auditory stimuli than in unfilled and visual ones. The principal evidence for a pacemaker speed interpretation is the slope effect noted in Experiments 2 and 3 and in Wearden et al. (1998, 2006), in which differences in estimates between F/U and auditory/visual stimuli are greater at longer durations than at shorter ones and change the slopes of regressions of estimates versus time reliably but change their intercepts much less frequently.

This finding is exactly the result predicted from conventional internal clock theory. On any trial, the number of pulses in the accumulator after a stimulus of some duration t is $rt + r(l_o - l_c)$, where r is pacemaker speed. The quantity $(l_o - l_c)$, the difference between the latency for opening the switch at the end of a stimulus (and stopping the accumulation of pulses), and the latency for closing it at the beginning of the stimulus (and starting the accumulation) may be negative, zero, or positive, depending on the relative values of l_o and l_c. This means that the latencies for closing the switch and opening it could have large absolute values yet produce a very small $(l_o - l_c)$ value if the two latencies were nearly identical. This suggests that although intercepts should be higher when pacemaker speed is higher, as r multiplies $(l_o - l_c)$, observation of the effect in data may be difficult and may be infrequently significant statistically, given the fact that $(l_o - l_c)$ is likely to be very small.

As noted earlier, research on the perception of unfilled intervals (sometimes contrasted with filled ones, although not always) has been very diverse, with a particularly noteworthy body of work being concerned with the perceptual effects of *markers* (i.e., the stimuli that begin and end the interval). Grondin (2003) reviews this work, much of which stems from his own laboratory. A complete discussion is beyond the scope of this article, but we mention just one effect that may link to our theoretical perspective, and this is the effect of *intermodal unfilled intervals*, that is, those that begin with one stimulus (e.g., an auditory stimulus) and end with another one (e.g., a visual stimulus). One particularly interesting result is that unfilled durations that are started and ended with auditory and visual stimuli are judged asymmetrically: The auditory–visual marker sequence generally produces longer judgments than does the visual–auditory one.

A potential theoretical explanation (discussed by Grondin, 2003) is based on an internal marker hypothesis. As Grondin (2003) himself notes, this hypothesis uses mechanisms rather similar to those proposed for the switch of the internal clock, and we translate it here in terms of the switch. Suppose that the latencies for closing and opening the switch differ for auditory and visual stimuli, with the switch operating more rapidly, in both cases, for auditory stimuli than for visual ones. Now, in an auditory–visual sequence, temporal accumulation begins rapidly but ends slowly, whereas the reverse is true for the visual–auditory sequence. This leads to more accumulation in the auditory–visual case than in the visual–auditory one and, thus, generally longer judgments, which are the results obtained. It may be that considerations of (a) pacemaker speed, (b) mean switch closing and opening latencies, and (c) switch closing and opening variance, taken together in a hybrid theory, can give a good prima facie explanation of many unfilled interval effects: Different pacemaker speeds account for overall differences in subjective duration between filled and unfilled intervals, mean switch latencies account for some marker effects, and switch variance may account for some perceived variability differences when judgments of duration of different stimulus types are carried out.

Although the notion that the filled stimuli seem, on average, longer than unfilled ones because of a faster pacemaker rate in the former case seems, to us, the simplest explanation of the results of the present experiments, it should be acknowledged that this is not the only possibility. Another sort of interpretation involves the amount of attention paid to stimuli of different types and the consequences of differences in allocation of attention on judgments of duration. Attentional effects on time judgments are well-known and reliable: For example, when people are required to perform an additional temporal or nontemporal task concurrently with a time judgment task, they behave as though the interval judged is shortened. Brown (1997) provides much experimental data and a good review of these and other attentional effects.

Although the effect of diverting attention away from timing by a concurrent task almost always produces a shortening of time judgments, the mechanism by which this manipulation operates is much less clear. Burle and Casini (2001) attempted to dissociate activation effects resulting from click trains such as those used in Experiment 3 of the present article, with attentional effects resulting from the presence or absence of a concurrent task. In their procedure, people were trained to produce a 1,100-ms time interval with their right hand and, on some production trials, also had to perform a concurrent choice reaction time task using two fingers of their left hand. Click trains were delivered during the productions, but these click trains were of either high or low intensity, with the low intensity clicks being adjusted so that they did not produce any activation effect. The louder clicks made the intervals produced shorter (cf. Penton-Voak et al., 1996), as did the presence of the reaction time task, but there was no statistical interaction between the two manipulations, suggesting an all-or-none effect of the concurrent task such as that which might occur with a single interruption in timing.

The fact that, at first sight, the filled interval illusion in our Experiments 2 and 3 manifests itself very clearly as a difference in slope appears to rule out a simple all-or-none difference in allocation of attention between filled and unfilled intervals. However, some attentional explanations offer the potential for explaining slope effects such as those found in Experiments 2 and 3 without appealing to a simple difference in pacemaker speed between filled and unfilled intervals. One such explanation is that of a flickering switch. According to this interpretation, an individual must consistently maintain attention to a stimulus to keep the switch connecting the pacemaker and the accumulator closed, and, without such attention, the switch tends to open spontaneously, cutting the pacemaker–accumulator connection and, thus, reducing the number of pulses accumulated (see Penney et al., 2000, p. 1784, for an explanation of differences in duration judgments of auditory and visual stimuli in these terms). The pacemaker–accumulator connection may be broken repeatedly and reconnected (the flicker of the switch), so longer stimulus durations involve greater loss of pulses than do shorter ones, with the consequence that the difference in perceived duration between two conditions increases as the stimuli being judged become longer. This explanation offers the possibility of explaining slope effects without positing the need for a difference in pacemaker speed between the conditions compared. The flickering switch explanation may not be challenged by Burle and Casini's (2001) result because, in their study, only a single reaction time task was presented during some production trials; thus, only a single switch of attention was necessary, whereas the flickering switch account requires multiple attentional switches.

In the present case, this interpretation would propose that individuals find it more difficult to maintain attention to unfilled intervals than to filled ones, with more flicker in the former case. Obviously, this interpretation mimics putative pacemaker speed effects very closely, so it is difficult to distinguish from the interpretation that we favor. However, although people may have problems maintaining task attention when the stimuli presented to them last a long time (see Wearden, 2002, for an example of a task in which some participants appeared to have this difficulty), it might be considered implausible that their attentional resources are challenged by the brief stimuli such as those used in the present study, particularly when there is no concurrent task required. In fact, inspection of Figures 2, 3, and 4 shows that an F/U duration estimate difference is generally present even when stimulus durations are very short. For example, if we take the simplest F/U comparisons in Experiments 2 and 3, that is, those comparisons between tones and stimuli that started and ended with clicks, there were significantly longer estimates in the filled case when the filled and unfilled stimuli were 203 ms long, Experiment 2, $t(14) = 2.55$, $p < .05$, and when they were 348 ms long, Experiment 3, $t(19) = 4.90$, $p < .001$. Such an effect is consistent with a pacemaker speed effect, which, we expect, would be manifested at short durations (although it would be manifested more clearly at longer ones) but would obviously require the switch to flicker very rapidly to produce its proposed effects.

Another attentional explanation that is difficult to distinguish from an account in terms of pacemaker rate is the attentional gate model of Zakay and Block (1998; see also Lejeune, 1998, for a discussion). This model proposes that the pacemaker is connected to the accumulator by a switch (which operates, more or less, automatically at stimulus onset and offset) and an attentional gate. A better physical metaphor for the proposed gate might be a squeezable tube: The more the tube is squeezed, the lower the rate of flow through it, with the result that the flow can be continuously varied. Zakay and Block (1998) propose that the attention paid to time modulates the attentional gate, with the gate being wider with more attention, and that the effect is a potentially continuous modulation of flow from the pacemaker to the accumulator. This attentional gate mechanism can, obviously, produce slope effects such as those obtained in our Experiments 2 and 3 if the attentional gate was wider for filled intervals than for unfilled intervals, with the expectation that the effect would be manifested even at short durations, as we have found.

Neither the flickering switch approach nor the attentional gate model has been developed in sufficient quantitative detail for the precise enumeration of predictions. One question that arises for both approaches concerns predictions about variability. The flickering switch model has the potential advantage that it may derive mean perceived duration and perceived duration variability from a common process: More frequent flickers may not only reduce total accumulation but may also make the accumulation in some fixed time more variable from trial to trial. However, it is unclear whether this explanation would predict differences in standard deviation and/or differences in coefficient of variation when filled and unfilled intervals are compared. In fact, exact prediction probably depends on the rate and periodicity of the flicker, but no published quantitative model currently exists. Likewise, it is unclear whether the attentional gate model would link the mean and the variability of perceived duration or, in general, what its predictions about variability might be.

In conclusion, the data presented in this article extend previous observations about properties of simple forms of the filled-duration illusion and show how the illusory differences in subjective duration between different stimulus types may be explained, at least in large measure, by using the type of pacemaker–accumulator internal clock proposed by SET. We make no claims to have explained all of the different phenomena discussed historically under the heading of the filled-duration illusion because, as mentioned in the introductory paragraph, it seems likely that these phenomena are heterogeneous in type and have many different causes, but we hope to show, here, that some simple forms of this effect are more easily understood than previously supposed.

References

Adams, R. D. (1977). Intervening stimulus effects on category judgments of duration. *Perception and Psychophysics, 21,* 527–534.

Allan, L. G. (1979). The perception of time. *Perception and Psychophysics, 26,* 340–354.

Allan, L. G. (1992). The internal clock revisited. In F. Macar, V. Pouthas, & W. J. Friedman (Eds.), *Time, action, and cognition* (pp. 191–202). Dordrecht, the Netherlands: Kluwer Academic.

Allan, L. G. (1998). The influence of the scalar timing model on human timing research. *Behavioural Processes, 44,* 101–117.

Allan, L. G., & Gibbon, J. (1991). Human bisection at the geometric mean. *Learning and Motivation, 22,* 39–58.

Brown, S. W. (1995). Time, change, and motion: The effects of stimulus movement on temporal perception. *Perception & Psychophysics, 57,* 105–116.

Brown, S. W. (1997). Attentional resources in timing: Inteference effects in concurrent temporal and nontemporal processing. *Perception & Psychophysics, 59,* 1118–1140.

Buffardi, L. (1971). Factors affecting the filled-duration illusion in the auditory, tactual, and visual modalities. *Perception & Psychophysics, 10,* 292–294.

Burle, B., & Casini, L. (2001). Dissociation between activation and attention effects in time estimation: Implications for internal clock models. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 195–205.

Church, R. M., Meck, W. H., & Gibbon, J. (1994). Application of scalar timing theory to individual trials. *Journal of Experimental Psychology: Animal Behavior Processes, 20,* 135–155.

Droit-Volet, S., Tourret, S., & Wearden, J. (2004). Perception of the duration of auditory and visual stimuli in children and adults. *Quarterly Journal of Experimental Psychology, 57,* 797–818.

Foley, A. J., Michaluk, L. M., & Thomas, D. G. (2004). Pace alteration and estimation of time intervals. *Perceptual and Motor Skills, 98,* 291–298.

Gibbon, J., & Church, R. M. (1984). Sources of variance in an information processing theory of timing. In H. L. Roitblat, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 465–488). Hillsdale, NJ: Erlbaum.

Gibbon, J., Church, R. M., & Meck, W. (1984). Scalar timing in memory. *Annals of the New York Academy of Sciences, 423,* 52–77.

Goldstone, S., & Lhamon, W. T. (1974). Studies of auditory–visual differences in human time judgment. 1. Sounds are judged longer than lights. *Perceptual and Motor Skills, 39,* 63–82.

Grondin, S. (2003). Sensory modalities and temporal processing. In H. Helfrich (Ed.), *Time and mind II*: *Information-processing perspectives* (pp. 62–77). Göttingen, Germany: Hogrefe & Huber.

Grondin, S., Roussel, M.-E., Lamache, P.-L., Roy, M., & Oullet, B. (2004). The structure of sensory events and the accuracy of time judgements. *Perception, 34,* 45–58.

Jones, L. A., & Wearden, J. H. (2003). More is not necessarily better: Examining the nature of the temporal reference memory component in timing. *Quarterly Journal of Experimental Psychology, 56B,* 321–343.

Jones, L. A., & Wearden, J. H. (2004). Double standards: Memory loading in temporal reference memory. *Quarterly Journal of Experimental Psychology, 57B,* 55–77.

Lejeune, H. (1998). Switching or gating? The attentional challenge in cognitive models of psychological time. *Behavioural Processes, 44,* 127–145.

Lejeune, H., Ferrara, A., Simons, F., & Wearden, J. H. (1997). Adjusting to changes in the time of reinforcement: Peak interval transitions in rats.

Journal of Experimental Psychology: Animal Behavior Processes, 23, 311–331.

Lejeune, H., & Wearden, J. H. (1991). The comparative psychology of fixed-interval responding: Some quantitative analyses. *Learning and Motivation, 22,* 84–111.

Lejeune, H., & Wearden, J. H. (2006). Scalar properties in animal timing: Conformity and violations. *Quarterly Journal of Experimental Psychology, 59,* 1875–1908.

Miki, A., & Santi, A. (2005). The perception of empty and filled time intervals by pigeons. *Quarterly Journal of Experimental Psychology, 58B,* 31–45.

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

Penney, T. P., Gibbon, J., & Meck, W. H. (2000). Differential effects of auditory and visual signals on clock speed and memory processes. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1770–1787.

Penton-Voak, I. S., Edwards, H., Percival, A., & Wearden, J. H. (1996). Speeding up an internal clock in humans? Effects of click trains on subjective duration. *Journal of Experimental Psychology: Animal Behavior Processes, 22,* 307–320.

Rammsayer, T. H., & Leutner, D. (1996). Temporal discrimination as a function of marker duration. *Perception & Psychophysics, 58,* 1213–1223.

Rammsayer, T. H., & Lima, S. D. (1991). Duration discrimination of filled and empty auditory intervals: Cognitive and perceptual factors. *Perception & Psychophysics, 50,* 563–574.

Santi, A., Keough, D., Gagne, S., & van Rooyen, P. (in press). Differential effects of empty and filled intervals on duration estimation by pigeons: Tests of an attention-sharing explanation. *Behavioural Processes.*

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

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock." *Psychological Monographs, 77,* 1–31.

Treisman, M., Faulkner, A., Naish, P. L. N., & Brogan, D. (1990). The internal clock: Evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception, 19,* 705–748.

Wearden, J. H. (1991a). Do humans possess an internal clock with scalar timing properties? *Learning and Motivation, 22,* 59–83.

Wearden, J. H. (1991b). Human performance on an analogue of an interval bisection task. *Quarterly Journal of Experimental Psychology, 43B,* 59–81.

Wearden, J. H. (1992). Temporal generalization in humans. *Journal of Experimental Psychology: Animal Behavior Processes, 18,* 134–144.

Wearden, J. H. (1999). "Beyond the fields we know": Exploring and developing scalar timing theory. *Behavioural Processes, 45,* 3–21.

Wearden, J. H. (2002). Traveling in time: A time-left analogue for humans. *Journal of Experimental Psychology: Animal Behavior Processes, 28,* 200–208.

Wearden, J. H. (2003). Applying the scalar timing model to human time psychology: Progress and challenges. In H. Helfrich (Ed.), *Time and mind II*: *Information-processing perspectives* (pp. 21–39). Göttingen, Germany: Hogrefe & Huber.

Wearden, J. H. (2004). Decision processes in models of timing. *Acta Neurobiologiae Experimentalis, 64,* 303–317.

Wearden, J. H. (2006, April). *Mission (Almost) Impossible: Modelling the verbal estimation of duration.* Paper presented at the meeting of the Experimental Psychology Society, Birmingham, England.

Wearden, J. H., & Bray, S. (2001). Scalar timing without reference memory: Episodic temporal generalization and bisection in humans. *Quarterly Journal of Experimental Psychology, 54B,* 289–310.

Wearden, J. H., Denovan, L., Fakhri, M., & Haworth, R. (1997). Scalar timing in temporal generalization in humans with longer stimulus durations. *Journal of Experimental Psychology: Animal Behavior Processes, 23,* 502–511.

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": Application of a model of the internal clock in humans. *Quarterly Journal of Experimental Psychology, 51B,* 97–120.

Wearden, J. H., & McShane, B. (1988). Interval production as an analogue of the peak procedure: Evidence for similarity of human and animal timing processes. *Quarterly Journal of Experimental Psychology, 40B,* 363–375.

Wearden, J. H., Philpott, K., & Win, T. (1999). Speeding up and (. . . relatively. . .) slowing down an internal clock in humans. *Behavioural Processes, 46,* 63–73.

Wearden, J. H., Todd, N. P. M., & Jones, L. A. (2006). When do auditory/visual differences in duration judgements occur? *Quarterly Journal of Experimental Psychology, 59,* 1709–1724.

Wearden, J. H., Wearden, A. J., & Rabbitt, P. (1997). Age and IQ effects on stimulus and response timing. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 962–979.

Witherspoon, D., & Allan, L. G. (1985). Time judgments and the repetition effect in perceptual identification. *Memory and Cognition, 13,* 101–111.

Zakay, D., & Block, R. A. (1998). New perspectives on prospective time estimation. In V. DeKeyser, G. d'Ydewalle, & A. Vandierendonck (Eds.), *Time and the dynamic control of behavior* (pp. 129–141). Göttingen, Germany: Hogrefe and Huber.

Received November 28, 2005
Revision received October 21, 2006
Accepted October 30, 2006 ■