Can we use verbal estimation to dissect the internal clock? Differentiating the effects of pacemaker rate, switch latencies, and judgment processes.

William J. Matthews

University of Essex, UK

Word Count: 200 (Abstract), 4920 (Main Text)

Running head: Verbal estimation and the internal clock

Keywords: internal clock; pacemaker-accumulator; time perception; verbal estimation

Author note

Affiliation: William J. Matthews, Department of Psychology, University of Essex.

Correspondence address: William Matthews, Department of Psychology, University of Essex, Colchester, CO4 3SQ, United Kingdom, E-mail: will@essex.ac.uk

Abstract

Behavioural timing is frequently assumed to be based on the accumulation of pulses from a pacemaker. In humans, verbal estimation is often used to determine whether the effect of factors which influence subjective time become more pronounced at longer durations - that is, if they affect the slope of the judgment function, consistent with a change in the rate of the pacemaker. Here, participants judged blank intervals marked by two squares which either did or did not differ in size. In Experiment 1, a small change in marker size produced shorter temporal judgments than a large change. This effect was independent of objective duration and indicates that the slope changes seen in previous work are not an inevitable artefact of the verbal estimation procedure. However, Experiments 2 and 3 included conditions where the markers did not change size and established (a) that the effect of marker size depends on the other stimuli presented during the experiment, and (b) that slope effects occur even when they cannot possibly be due to a change in the rate of the pacemaker. Taken together, these results urge some caution in the use of verbal estimation as a methodology for deconstructing the putative internal clock.

1. Introduction

Many theories of behavioural timing assume the existence of a pacemaker-accumulator system (e.g., Allan, 1998; Gibbon et al. 1984; Treisman, 1963; Ulrich et al., 2006; Wearden, 1992; Zakay & Block, 1997; for contrasting accounts, see e.g., Machado, 1997; Mauk & Buonomano, 2004; Staddon & Higa, 1999; Wackermann & Ehm, 2006). There are various instantiations of this idea, but the core assumption is that timing is based on a pacemaker which emits pulses; at the start of the to-be-timed interval, a switch closes so that these pulses flow into an accumulator. At the end of the timed interval, the switch opens, cutting the flow, and the number of accumulated pulses forms the basis for time perception. The accumulated counts may be compared with previously encoded durations to provide the basis for judgment and action (for recent discussions, see Jones & Wearden, 2003; 2004; Ogden & Jones, 2009; Ogden, Wearden & Jones, 2008).

Despite the importance of accurate timing, many non-temporal factors affect subjective duration. For example, tone stimuli are judged longer than lights of equal duration (Goldstone & Lhamon, 1874), continuous stimuli are judged longer than empty intervals defined by brief markers (e.g., Wearden et al., 2007), and stimuli which stand out from the background have longer subjective durations than those which do not (e.g., Matthews et al., 2010). Within the pacemaker-accumulator framework, such effects may be due to a change in switch latencies, a change in pacemaker rate, or both. If a manipulation influences the switch latencies (that is, the time taken to begin or end the timing process) then the effect will be independent of the objective duration of the stimulus. For example, suppose that the time taken to open the switch and terminate pulse accumulation at the end of the stimulus is longer in one condition than another. The number of accumulated counts will be greater and the subjective duration longer, and this effect will be independent of the physical duration of the stimulus. By contrast, if the pacemaker runs faster in one condition than another then the difference between them will increase as the physical duration is extended because the number of accumulated counts is obtained by multiplying the rate of the pacemaker and the length of the accumulation period (Penton-Voak et al., 1996).

One way to deconstruct the influence of experimental manipulations on the operation of the putative pacemaker-accumulator system in humans is to present stimuli of varying duration under two or more conditions, and to elicit numerical estimates of their duration. These estimates are then plotted against physical time; if the lines for the two conditions diverge as objective time increases, it is taken as evidence that the manipulation affects the rate of the pacemaker.

Figure 1 shows the results of several published experiments which have used this method. The top left panel shows a comparison of tone and light stimuli (Wearden et al., 1998); the top right panel compares filled and empty intervals (Wearden et al., 1997); the bottom left panel compares stimuli which have or have not been preceded by a rapid series of clicks designed to increase the rate of the pacemaker (Penton-Voak et al., 1996); the bottom right panel compares quiet and loud tones presented in silence (Matthews et al., 2010). In every case, verbal estimates increase with objective duration, but the slope is greater for one condition than the other. These papers have therefore concluded that stimulus modality, filling the interval, changing stimulus intensity, and playing a click train all influence the rate of the pacemaker [1].

The current paper is concerned with the use of the verbal estimation methodology as a tool for dissecting the effect of experimental manipulations on the operation of the putative internal clock. The starting point for the these experiments was the idea that differences in pacemaker rate are not the only explanation for the slope effects seen in Figure 1. It is conspicuous that *every* study using the verbal estimation paradigm has found a slope effect; none has found a pure intercept effect. This raises the question of whether the verbal estimation procedure is capable of producing such an effect and, correspondingly, of whether the slope effects might be an artefact of the process by which subjective duration is converted into a numerical estimate. In the broadest terms, it may be that *any* manipulation which affects verbal estimates does so in a way which is multiplicative with the duration of the stimulus.

The first experiment sought to establish whether the verbal estimation paradigm is able to produce an intercept effect in a situation in which a change in pacemaker rate is impossible. Participants timed a blank interval defined by two markers – squares which differed in size. The first marker was the same on every trial. On some trials, the second marker was a slightly different size from the first (a small jump); on other trials, the second marker was a very different size from the first (a large jump). Xuan et al. (2007) found that large changes in marker size increased the subjective duration of the interval between the markers (although they did not examine the effects of varying stimulus duration). Within the pacemaker-accumulator framework, this effect can only be due to a change in the latency with which the switch is opened at the end of timing; it cannot be due to a change in pacemaker rate because the first marker and the display during the to-be-timed interval are identical on every trial. The verbal estimation method should therefore produce a change in intercept but no slope effect. However, if the slope effect found in previous studies arises from a judgment heuristic wherein the effects of experimental manipulations on verbal estimates are multiplicative with raw subjective time, the same kind of slope effect will be found here. The presence of a slope effect in the current experiment would therefore question the pacemaker-rate explanation for the effects of modality, intensity, click-trains and filled intervals shown in Figure 1.

2. Experiment 1

2.1. Method

2.1.1. Participants. Twenty four participants (10 male, aged 19-36, *M* = 25.3 years, *SD* = 4.9 years) took part for a payment of £4. Two additional participants were discarded for failing to follow the instructions.

2.1.2. Stimuli. The stimuli were black empty squares (drawn with lines 5 pixels thick) shown against a white background on a 19" CRT monitor(1024x768 pixels, 85 Hz) viewed from approximately 90 cm through the window of a sound-attenuating chamber. There were four sizes of square, similar to those used by Xuan et al. (2007). The sizes were 40x40 pixels (c. 0.9° visual angle), 60x60 pixels (1.3°), 150x150 pixels (3.3°), and 170x170 pixels (3.7°). For convenience, these will be numbered 1-4 in order of increasing size.

2.1.3. Design and Procedure. There were three independent variables: the duration of the timed interval, the size of the first marker, and the magnitude of the jump in marker size between the first and second markers. For half of the participants the first marker was always square 1 (the smallest square) and the second marker was square 2 on half of the trials (Small Jump) and square 4 on the other half (Big Jump). For the remaining participants, the first marker was always square 4 and the second marker was square 3 on half of the trials (Small Jump) and square 1 on the other half (Big Jump). Thus, for both groups of participants, there was a small jump in marker size on some trials and a big jump on others.

The sequence of events on each trial was: blank interval for 1 s, first marker for 259 ms; blank to-be-timed interval for 71, 176, 282, 388, 494 or 600 ms; second marker for 259 ms; blank response window. (The to-be-timed intervals were chosen to cover the range of durations where the slope effects were most pronounced in previous work.) Presentation was controlled by DMDX (Forster & Forster, 2003). After offset of the second marker, participants typed their estimate of the length of the blank interval, in milliseconds. They were reminded that 1 s = 1000 ms and told that the interval would never be shorter than 50 ms or longer than 1000 ms. The digits that the participant typed were not visible on the screen, but if they made a mistake they could press a key to clear their entry and re-type the response from scratch. After typing their judgment they pressed the Return key to trigger the next trial. Participants completed 5 blocks of 24 trials with two presentations of each duration-jump size combination per block, in random order. In all experiments, a handful of trials were discarded because of display timing errors or because responses were missing/outside the range 50-1000 ms (1.7% of trials in Experiment 1; 1.3% in Experiments 2 and 3).

2.2. Results and Discussion

The data are shown in Figure 2. Panel (a) shows the results from those participants for whom the first marker was always the smallest square; panel (b) shows the results when the first marker was always the largest square; panel (c) shows the data collapsed across these two groups. Inspection suggests that when there is a big jump in marker size the interval is judged longer than when there is a small jump in marker size. This effect appears to be independent of both the objective duration and the size of the first marker.

A 2 (size of first marker) x 2 (jump size) x 6 (duration) mixed ANOVA confirmed these impressions. As one would expect, judgments were larger for longer durations, $F(1.54, 33.84) = 62.46$, $p < .001$, $\eta_p^2 = .74$ (here and at certain points below a Huynh-Feldt correction was applied because of violations of sphericity). Intervals defined by a big jump in marker size were judged longer than those defined by a small jump, $F(1,23) = 20.27$, $p < .001$, $\eta_p^2 = .48$, and this effect was independent of duration, $F(4.45, 97.85) = 0.93$, $p = .458$, $\eta_p^2 = .04$. None of the effects involving the size of the first marker were significant (all *Fs* < 1).

Following the example of previous work, an additional analysis was conducted in which separate regression lines were fitted for each participant (see e.g., Wearden et al., 1998). The mean intercept and slope coefficients are shown in Table 1. A 2 (first marker) x 2 (jump size) mixed ANOVA established that the intercepts were larger in the big jump condition than in the small jump condition, $F(1,22) = 11.22$, $p = .003$, $\eta_p^2 = .34$, but were unaffected by the size of the first marker (both main effect and interaction *F*s < 1). The slope coefficients were unaffected by jump size, $F(1,22) = 1.85$, $p = .187$, $\eta_p^2 = .08$, and did not depend on the size of the first marker (both main effect and interaction *F*s < 1). This pattern mirrors that from the basic ANOVA, above.

Thus, when participants time a blank interval defined by two markers, a large change in marker size produces an increase in estimated duration. Crucially, this effect is additive with physical time; there is no indication of the type of slope effect seen for other manipulations. This is unlikely to be due to lack of statistical power: the power to detect a duration*jump size interaction with an effect size of $\eta_p^2 = .1$ (smaller than the smallest effect recently reported by Matthews et al., 2010) is approximately 99% (Erdfelder et al., 1996).

This is the first verbal estimation experiment to find a manipulation which produces an intercept effect, and it serves as proof that the paradigm is capable of detecting a change in intercept of the type predicted by a switch effect in the pacemaker-accumulator framework. This is reassuring: had the current experiment produced the same pattern of results as those of Penton-Voak et al. (1996), Wearden et al. (1998), Matthews et al., (2010), and others, it would strengthen the argument that these slope effects are the result of a judgment heuristic or some other artefact of the verbal estimation procedure, rather than a pacemaker effect.

Within the pacemaker-accumulator framework, the current results are best explained as a switch effect. When the difference between the second marker and the first is more pronounced, judged duration increases, presumably because the time taken to process the second marker (and hence to open the switch) is longer. It is noticeable that it is the difference between the two markers, not absolute marker size, which determines this effect: based on the evidence of this experiment, switch latencies seem to depend on relative stimulus magnitudes rather than absolute magnitudes.

Although the results of Experiment 1 indicate that the verbal estimation paradigm is capable of producing an intercept effect in a situation for which a difference in pacemaker rate is impossible, Experiments 2 and 3 raise more doubts about the use of the verbal estimation procedure. In Experiment 1 the second marker was always a different size from the first; Experiments 2 and 3 examine what happens when the first and second markers are the same size.

3. Experiment 2

Experiment 2 was similar to Experiment 1, except that the to-be-timed intervals were defined by markers which were either both small or both large.

3.1. Method

3.1.1. Participants. Twenty four participants (16 male, aged 19-39, M = 24.8 years, SD = 4.3 years) took part for a payment of £4. None had participated in Experiment 1. Two additional participants were discarded for failing to follow instructions.

3.1.2. Stimuli. The smallest and largest squares from Experiment 1 were used. For convenience, they will be referred to as “small” and “large”, respectively.

3.1.3. Design and Procedure. The two independent variables were duration and marker size; both were manipulated within participants. On half of the trials the small square was used as both the first and second marker (condition Small-Small); on the other half, the large marker was used as both the first and second marker (condition Large-Large). Participants completed 5 blocks of 24 trials with each duration-marker size combination occurring twice in each block. In other respects the experiment was identical to Experiment 1.

3.2. Results and Discussion

The results are shown in Figure 3. A 2x6 repeated measures ANOVA revealed that, as one would expect, judgments increased with increasing duration, $F(1.69, 38.91) = 138.47$, $p < .001$, $\eta_p^2 = .86$. However, there was no indication of a main effect of marker size and no interaction between size and duration, both *F*s < 1. Fitting regression lines to individual participant data revealed that the intercepts were the same for the Small-Small ($M = 58.94$, $SD = 102.5$) and Large-Large conditions ($M = 52.74$, $D = 98.92$), $t < 1$. Similarly, the slopes were unaffected by marker size (for Small-Small, $M = 0.93$, $SD = 0.35$; for Large-Large, $M = 0.97$, $SD = 0.36$; $t < 1$).

Experiment 2 therefore found no evidence of a difference between the Small-Small and Large-Large conditions: when the first and second markers are the same size, the judged duration does not depend on the absolute size of the markers. Taken in isolation, this result is of limited interest – although it suggests that it was indeed the *change* in size between the first and second markers which was crucial to the intercept effect in Experiment 1. However, the failure to find an effect in this experiment will take on a different complexion following Experiment 3.

4. Experiment 3

Experiment 3 was similar to Experiment 2 except for the addition of two conditions in which the second marker was different from the first.

4.1. Method

4.1.1. Participants. Twenty four participants (13 male, aged 21-49, $M = 26.5$ years, $SD = 47.$ years) took part for a payment of £4 (one declined payment). None had participated in Experiments 1 or 2. Two additional participants were discarded for failing to follow instructions.

4.1.2. Stimuli. The small and large squares from Experiment 2 were used.

4.1.3. Design and Procedure. Three independent variables were manipulated within participants: duration, size of the first marker, and size of the second marker. The size of the first and second marker were factorially combined to give four conditions: a small first marker followed by a small second marker (Small-Small); a small first marker followed by a large second marker (Small-Large); a large first marker followed by a small second marker (Large-Small); and a large first marker followed by a large second marker (Large-Large). Participants completed 9 blocks of 24 trials, with each marker condition – duration combination occurring once per block, in random order. In other respects the procedure was like Experiments 1 and 2.

4.2. Results and Discussion

The results are shown in Figure 4. An initial 2 x 2 x 6 within-subjects ANOVA established a significant three-way interaction between first marker, second marker, and duration, $F(3.33, 76.57) = 4.60$, $p = .004$, $\eta_p^2 = .17$. This interaction was decomposed by separate two-way ANOVAs.

When the first marker was small, the intervals were judged longer when the marker changed size than when it stayed the same, $F(1,23) = 44.73$, $p < .001$, $\eta_p^2 = .66$, but this effect diminished as the to-be-judged interval increased, $F(3.99,91.82) = 3.53$, $p = .01$, $\eta_p^2 = .13$. The same pattern was found when the first marker was large [for the main effect of a change in marker size $F(1,23) = 11.45$, $p = .003$, $\eta_p^2 = .33$; for the interaction with duration: $F(3.33, 76.67) = 4.44$, $p = .005$, $\eta_p^2 = .16$]. In addition, judgments were bigger in the Large-Large condition than the Small-Small condition, $F(1,23) = 17.89$, $p < .001$, $\eta_p^2 = .44$, an effect which was independent of duration, $F(3.77, 86.70) = 2.42$, $p = .058$, $\eta_p^2 = .10$. As one would expect, all of the foregoing analyses revealed significant effects of duration, with larger judgments for longer stimuli.

Separate regression coefficients were estimated for each participant and are shown in Table 2. The intercepts are larger when the first and second marker were different than when they were the same: a 2x2 repeated measures ANOVA established the significance of this interaction, $F(1,23) = 49.10$, $p < .001$, $\eta_p^2 = .68$. Conversely, the slopes are steeper when the first and second markers were the same size than when they were different, $F(1,23) = 23.12$, $p < .001$, $\eta_p^2 = .50$. This pattern of results accords with the ANOVAs reported above.

To summarize: the judgment function has a flatter slope and a larger intercept when the interval is defined by markers of different sizes than when the first and second marker are identical. This is true irrespective of whether the first marker is large or small. Furthermore, in contrast to Experiment 2, there was a difference between the Small-Small and Large-Large conditions. These results have interesting implications for the use of the verbal estimation procedure, which are discussed below.

5. General Discussion

In Experiment 1, observers judged blank intervals defined by two squares; the second marker was always a different size from the first, and big jumps in marker size produced larger duration estimates than small jumps. Crucially, this effect was independent of the to-be-judged interval. In Experiment 2, the to-be-timed intervals were defined by two squares which were either both small or both large. Marker size had no effect on judgment. In Experiment 3, the first and second markers were either identical or different sizes. Changes in marker size elicited larger judgments, and this effect diminished as the stimulus duration increased. In addition, when both markers were large the estimated duration was longer than when both markers were small, in contradistinction to the results of Experiment 2. These results cast new light on the use of verbal estimation as a technique for dissecting the putative internal clock.

Consider first Experiment 1. This study is the first verbal estimation experiment to find a manipulation which produces a change in the intercept of the judgment function but no change in its slope. For internal clock models, this is the pattern that one would expect from a change in switch latencies when there is no change in the rate of the pacemaker. Of course, this does not necessarily mean that the results of Experiment 1 are due to the relative size of the second marker affecting the time taken to open the switch and end the accumulation of pulses: various accounts of behavioural timing which do not assume a pacemaker-accumulator system at all might equally well explain this pattern (e.g., Machado, 1997; Staddon & Higa, 1999). However, the key point is that Experiment 1 demonstrates that the verbal estimation procedure is capable of producing an intercept effect consistent with a change in switch latencies. It is not the case that any manipulation which affects subjective duration produces a slope effect as a procedural artefact.

Experiments 2 and 3 are more problematic, for two reasons. The first issue concerns a discrepancy in the results. Experiment 2 found no difference between Small-Small and Large-Large conditions, whereas Experiment 3 found a highly significant tendency for Large-Large intervals to be judged longer than Small-Small ones. The instructions, stimuli and procedure for these experiments were identical, and they were conducted using the same participant pool (but different participants) within a few days of each other; the main difference was that in Experiment 3, the Small-Small and Large-Large conditions were intermixed with Small-Large and Large-Small trials. (An additional difference is that a research assistant ran Experiment2, perhaps introducing subtle changes in the interaction with participants.) It is unlikely that the null result in Experiment 2 is due to low power: the power to detect an effect like that seen in Experiment 3 is approximately 99% (Erdfelder et al., 1996). It therefore seems that the effects of a particular manipulation depend upon the other items that are presented for judgement during the session. This kind of context effect is common in studies of perceptual discrimination and identification in both humans and animals (e.g., Hinson & Lockhead, 1986; Matthews & Stewart, 2008; Schneider & Parker, 1990), and the present results emphasize that any conclusions drawn from the verbal estimation procedure are conditional upon the specific experimental context: manipulations do not have absolute effects on the switch latencies or pacemaker rate of an internal clock. Rather, these effects depend upon the other items included in the session. A similar conclusion is urged by data from Matthews et al. (2010), who found that the background against which stimuli were presented influenced their effects on the pacemaker rate. The most likely explanation is that judgments of a given stimulus are partly based on comparisons with the other items in the experimental ensemble. For example, Brown et al (2005) found that temporal bisection curves depend on the skew of the distribution of to-be-judged intervals. Similarly, comparison of the non-temporal properties of the items in the stimulus set also shapes judgments of duration. For example, Gomez and Robertson (1979) found that large stimuli were judged to last longer than small ones, but only when observers encountered both sizes in the experimental session. It seems that participants faced with difficult temporal discriminations fall back on more straightforward relative judgments and the heuristic use of non-temporal information when producing their responses.

The second, more serious issue raised by Experiment 3 concerns the use of slope and intercept effects as indicators of switch/pacemaker effects in an internal clock model. In this experiment, a change in marker size led to an increase in intercept but a decrease in slope. It is worth reiterating the logic behind the use of the verbal estimation procedure. If a manipulation influences the latency with which timing is begun or ended then the effect will be independent of the length of the to-be-timed interval. By contrast, if a manipulation influences the rate of the pacemaker during the timed interval then its effect will become more pronounced as the timed duration increases, leading to a slope effect. Now consider the difference between the Small-Small and Small-Large conditions in Experiment 3. The "start timing" signal (the offset of the first square) and the timed interval (the blank display between the squares) is identical in both cases – only the nature of the "stop timing" signal differs. The rate of the internal clock during the to-be-timed interval must be the same in both conditions because the subject observes exactly the same sequence of events. Thus, Experiment 3 produced a slope effect in a situation where this cannot possibly be due to a difference in the rate of the pacemaker during the timed interval.

The fact that slope changes can arise for some reason other than a change in the rate of the pacemaker means that we must be somewhat cautious about using verbal estimation as a tool for dissecting the putative internal clock. As noted above, slope effects have been taken as evidence that various manipulations – including modality, intensity, and playing a click train prior to the stimulus - influence the rate of the pacemaker. It would be premature to reject these conclusions, but the current data raise the possibility that an alternative process is at work.

An obvious question is: if the slope changes in Experiment 3 are not due to the pacemaker, what are they due to? One possibility is that they arise from the use of a bounded judgment scale. Participants were told that the durations would be in the range 50 to 1000 ms; if the longest subjective durations for both stimulus categories were outside this range, a ceiling effect might arise such that the judgments for the different conditions are pushed together at the top of the range, creating a difference in the slope of the judgment function. This is unlikely, however, as the mean judgments for the longest stimuli were well below the maximum value of 1000 ms.

A more plausible suggestion is that the effect is due to an interaction between the processing of the two markers. That is, the time taken to register the second marker depends on its relationship to the first, and the effects of this interaction change over time. One mechanism by which this might occur involves repetition priming. When the same stimulus occurs twice in succession, the processing of the second occurrence is usually faster and more efficient than that of the first (Henson & Rugg, 2003; for discussions of this phenomenon in the context of time perception, see Pariyadath & Eagleman, 2008, and Eagleman & Pariyadath, 2009). Thus, in the Small-Small and Large-Large conditions, the processing of the second marker is likely to be quicker than in the Small-Large and Large-Small conditions. From the perspective of internal clock theories, more efficient processing of the second marker corresponds to a reduction in the time taken to stop the flow of pulses into the accumulator, with a corresponding decrease in subjective duration. If repetition priming occurs, and if the extent of the priming diminishes as the interval between the first and second presentations increases, then the effect of the end marker manipulation will decrease as the to-be-timed interval increases – the pattern seen in Experiment 3.

It will be important to establish whether the slope effects found in previous studies are also due to whatever process produced the effects seen in Experiment 3. There are at least two reasons for suspecting that they are not. First, Experiments 1 and 2 show that the kind of interaction found in Experiment 3 is not ubiquitous. Second, the slope effect in Experiment 3 is rather different from that in previous work in that there was also an intercept shift (compare Figures 1 and 4). However, with existing data there is no way of knowing whether or not the process underlying the slope effects in Experiment 3 was also at work in other studies. It is possible that the results of previous verbal estimation experiments arose from a combination of pacemaker effects, switch latency effects, and marker interactions, rather than pure pacemaker rate changes. The best interim conclusion is that we should be circumspect about the use of verbal estimation as a method for establishing whether a manipulation which influences temporal judgments acts on the switch or pacemaker components of the putative internal clock.

To summarize: these experiments make four contributions. First, they provide further evidence that the timing of empty intervals depends on the markers that define those intervals (e.g., Grondin, 1993; Grondin et al., 1996), and in particular that changes in the size of the marker produce substantial effects on temporal judgments (Ono & Kitazawa, 2009; Xuan et al., 2007). Second, and more relevant to the main focus of this paper, Experiment 1 shows that the verbal estimation procedure can produce an intercept effect of the type predicted when a manipulation affects the switch latencies of an internal pacemaker-accumulator system. The possibility of obtaining this effect is central to the use of verbal estimation as a tool for dissecting the internal clock, but the pattern has never previously been reported. Third, the comparison between the results of Experiments 2 and 3 suggest that whether a manipulation affects the putative pacemaker-accumulator system or not depends upon the other stimuli that are presented during the session. This urges some caution about treating the conclusions of verbal estimation tasks (or perhaps any time judgment experiments) as concrete indicators of how a particular manipulation influences the operation of the putative internal clock. Fourth, and perhaps most important, the results of Experiment 3 suggest that slope effects can emerge even when a change in the rate of the pacemaker is impossible. It is not clear exactly what underlies these effects, but they urge caution in the use of the verbal estimation procedure.

Footnotes

[1] Penney, Gibbon, and Meck (2000) offer an alternative account of the difference between visual and auditory stimuli, wherein the switch between pacemaker and accumulator “flickers” open and closed at a rate that depends on the attention paid to the stimulus. The predictions of this account are very similar to those resulting from a change in the rate of the pacemaker (Wearden et al., 2007), and both theories posit that slope effects arise because the difference in the number of accumulated pulses increases with increasing physical duration.

Acknowledgments

This work was supported by a grant from the University of Essex Research Promotion Fund. I thank Rebeccah Billing for help with data collection, and Neil Stewart and two anonymous referees for helpful comments on an earlier version.

References

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

Brown, G.D.A., McCormack, T., Smith, M., & Stewart, N. (2005). Identification and bisection of temporal durations and tone frequencies: Common models for temporal and nontemporal stimuli. J. Exp. Psychol. Hum. Percept. Perform. 31, 919-938.

Eagleman, D.M., & Pariyadath, V. (2009). Is subjective duration a signature of coding efficiency? Philos. Trans. R. Soc. Lond. B. 364, 1841-1851.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. Behav. Res. Methods Instrum. Comput. 28, 1-11.

Forster, K.I., & Forster, J.C. (2003). DMDX: A Windows display program with millisecond accuracy. Behav. Res. Methods Instrum. Comput. 35, 116-124.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. Allan (Eds.), *Annals of the New York Academy of Sciences, Volume 423: Timing and time perception*. New York: New York Academy of Sciences.

Goldstone, S., & Lhamon, W. T. (1974). Studies of auditory-visual differences in human time judgments: 1. Sounds are judged longer than lights. Percept. Mot. Skills 39, 63-82.

Gomez, L.M., & Robertson, L.C. (1979). The filled-duration illusion: The function of temporal and nontemporal set. Percept. Psychophys. 25, 432-438.

Grondin, S. (1993). Duration discrimination of empty and filled intervals marked by auditory and visual signals. Percept. Psychophys. 54, 383-394.

Grondin, S., Ivry, R.B., Franz, E., Perreault, L., & Metthé, L. (1996). Markers' influence on the duration discrimination of intermodal intervals. Percept. Psychophys. 58, 424-433.

Henson, R.N.A., & Rugg, M.D. (2003). Neural response suppression, haemodynamic repetition effects, and behavioural priming. Neuropsychologia 41, 263-270.

Hinson, J.M., & Lockhead, G.R. (1986). Range effects in successive discrimination. J. Exp. Psych. Anim. Behav. Process. 12, 270-276

Jones, L.A., & Wearden, J.H. (2003). More is not necessarily better: Examining the nature of the temporal reference memory component in timing. Q. J. Exp. Psychol. 56B, 321-343.

Jones, L.A., & Wearden, J.H. (2004). Double standards: Memory loading in temporal reference memory. Q. J. Exp. Psychol. 57B, 55-77.

Machado, A. (1997). Learning the temporal dynamics of behaviour. Psychol Rev. 104, 241-265.

Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. Can. J. Exp. Psychol. 57, 203-220.

Matthews, W.J., & Stewart, N. (2008). The effect of stimulus range on two-interval frequency discrimination. J. Acoust. Soc. Am. 123, EL45-EL51.

Matthews, W. J., Stewart, N., & Wearden, J. H. (2010). Stimulus intensity and the perception of duration. J. Exp. Psychol. Hum. Percept. Perform. In Press.

Mauk, M.D., & Buonomano, D.V. (2004). The neural basis of temporal processing. Annu. Rev. Neurosci. 27, 307-340.

Ogden, R.S., & Jones, L.A. (2009). More is still not better: Testing the perturbation model of temporal reference memory across different modalities and tasks. Q. J. Exp. Psychol. 62, 909-924.

Ogden, R. S., Wearden, J. H., & Jones, L. A. (2008). The Remembrance of Times Past: Interference in Temporal Reference Memory. J. Exp. Psychol. Hum. Percept. Perform. 26, 1770-1787.

Ono, F., & Kitazawa, S. (2009). The effect of marker size on the perception of an empty interval. Psychon. Bull. Rev. 16, 182-189.

Pariyadath, V., & Eagleman, D.M. (2008). Brief subjective durations contract with repetition. J. Vis. 8, 1-6

Penney, T.B., Gibbon, J., & Meck, W.H. (2000). Differential effects of auditory and visual signals on clock speed and temporal memory. J. Exp. Psychol. Hum. Percept. Perform. 34, 1524-1544.

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. J. Exp. Psychol. Anim. Behav. Process. 22, 307-320.

Schneider, B., & Parker, S. (1990). Does stimulus context affect loudness or only loudness judgments? Percept. Psychophys. 48, 409-418.

Staddon, J.E.R., & Higa, J.J. (1999). Time and memory: Towards a pacemaker-free theory of interval timing. J. Exp. Anal. Behav. 71, 215-251.

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock". Psychol. Monogr. 77, 1-31.

Ulrich, R., Nitschke, J., & Rammsayer, T. (2006). Crossmodal temporal discrimination: Assessing the predictions of a general pacemaker-counter model. Percept. Psychophys. 68, 1140-1152.

Wackermann, J., & Ehm, W. (2006). The dual klepsydra model of internal time representation and time reproduction. J. Theor. Biol. 239, 482-493.

Wearden, J. H. (1992). Temporal generalization in humans. J. Exp. Psychol. Anim. Behav. Process. 18, 134-144.

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "Sounds are judged longer then lights": Application of a model of the internal clock in humans. Q. J. Exp. Psychol. 51B, 97-120.

Wearden, J. H., Norton, R., Martin, S., & Montford-Bebb, O. (2007). Internal clock processes and the filled-duration illusion. J. Exp. Psychol. Hum. Percept. Perform. 33, 716-729.

Xuan, B., Zhang, D., He, S., & Chen, X. (2007). Larger stimuli are judged to last longer. J. Vis. 7, 1-5.

Zakay, D., & Block, R. A. (1997). Temporal cognition. Curr. Dir. Psycho. Sci. 6, 12-16.

Table 1. Mean intercept and slope coefficients from Experiment 1 (standard deviations in parentheses).

	Small first marker		Large first marker	
	Small Jump	Big Jump	Small Jump	Big Jump
Intercept	116.62	187.94	119.27	183.90
	(143.33)	(194.84)	(177.64)	(181.29)
Slope	0.65	0.72	0.71	0.72
	(0.41)	(0.45)	(0.39)	(0.36)

Table 2. Mean intercept and slope coefficients for the four conditions of Experiment 3 (standard deviations in parentheses).

	Small-Small	Small-Large	Large-Small	Large-Large
Intercept	30.12	129.04	112.51	25.54
	(70.62)	(98.98)	(103.75)	(68.54)
Slope	0.71	0.61	0.64	0.79
	(0.32)	(0.29)	(0.36)	(0.32)

Figure 1. Data from four experiments demonstrating a slope (pacemaker) effect in temporal estimation. (a) Data from Wearden et al. (1998) comparing the judged duration of tones and lights. (b) Data from Wearden et al. (2007) comparing the judged duration of filled intervals (continuous 500 Hz tones) and empty intervals (silent intervals flanked by 10 ms 1000 Hz clicks). (c) Data from Penton-Voak et al. (1996) comparing the judged durations of a blue square preceded/not preceded by a 5 Hz train of clicks for 5 s. (d) Data from Matthews et al. (2010) comparing the judged duration of quiet (59 dBA) and loud (80 dBA) 500 Hz tones presented in silence. In every case, the effect of the stimulus manipulation increases as the physical duration lengthens.

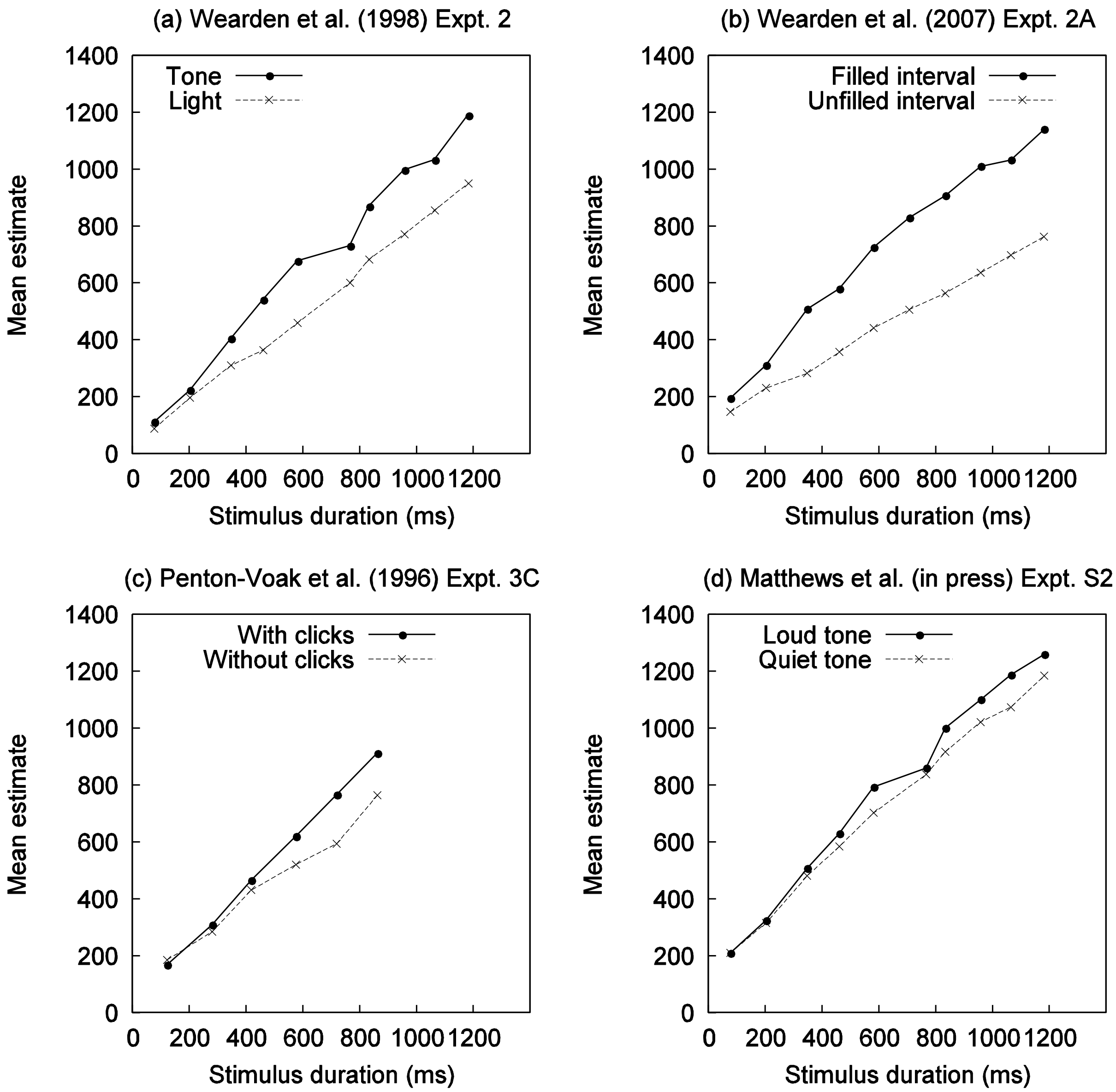

Figure 2. Results of Experiment 1. Panel (a) shows the results when the first marker was small and the second marker was larger; panel (b) shows the results when the first marker was large and the second marker was smaller. These two groups did not differ, and the collapsed data are shown in panel (c). There is a clear intercept effect, but no indication of a slope effect. The error bars show the standard error of the mean, calculated separately for each data point. Note that for a within-subject such as this one, these error bars provide no indication of the significance of differences between means (e.g., Masson & Loftus, 2003). (Violations of sphericity meant that it was not possible to use a pooled error term.)

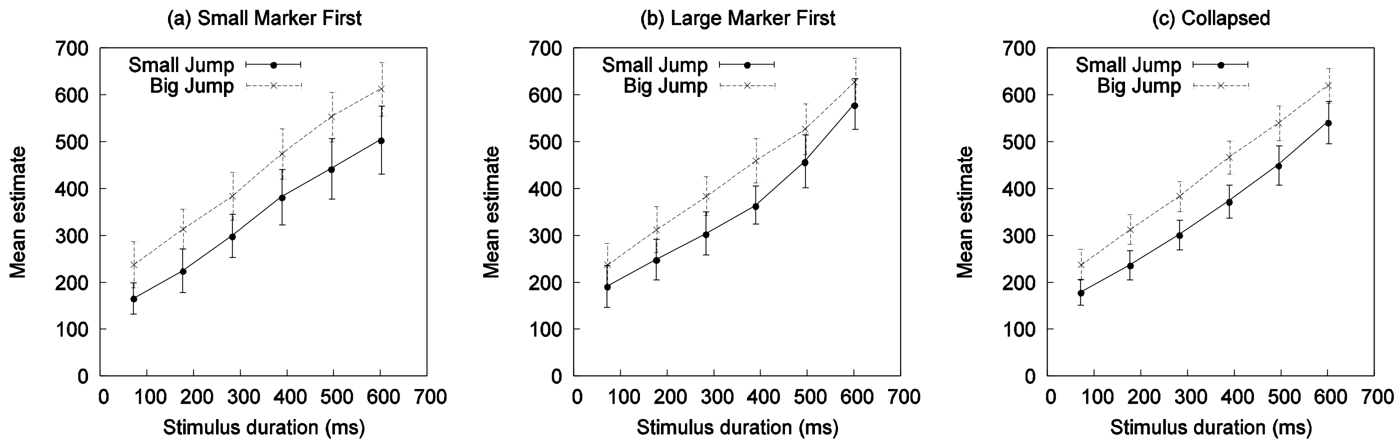

Figure 3. Results of Experiment 2.

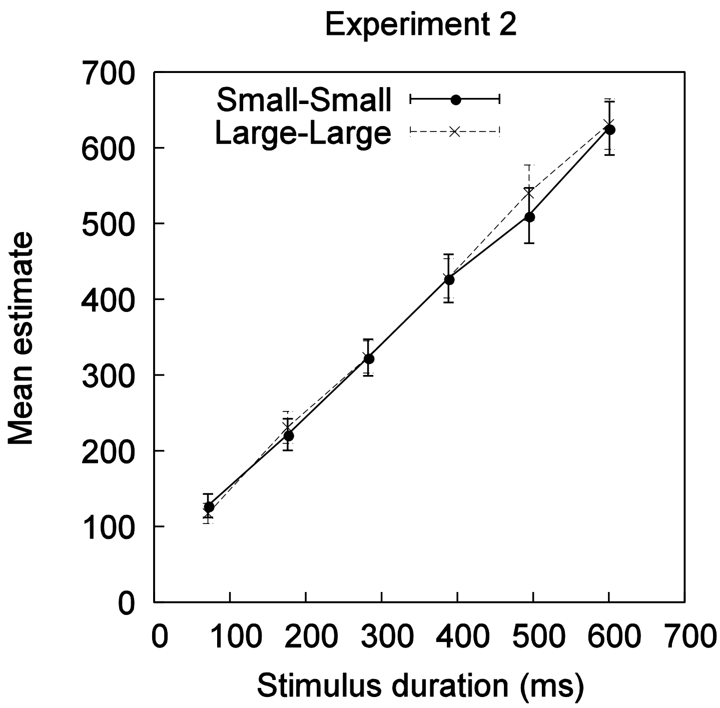

Figure 4. Results of Experiment 3. The left panel shows the results when the first marker was the small square; the right panel shows the results when the first marker was the large square. In both cases, the judged duration is longer when the second marker is different from the first, and this effect diminishes at longer durations.