THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 1998, *51B* (2), 97–120

# Why "Sounds Are Judged Longer Than Lights": Application of a Model of the Internal Clock in Humans

J. H. Wearden, H. Edwards, M. Fakhri, A. Percival

*University of Manchester, Manchester, U.K.*

Three experiments, using temporal generalization and verbal estimation methods, studied judgements of duration of auditory (500-Hz tone) and visual (14-cm blue square) stimuli. With both methods, auditory stimuli were judged longer, and less variable, than visual ones. The verbal estimation experiments used stimuli from 77 to 1183 msec in length, and the slope of the function relating mean estimate to real length differed between modalities (but the intercept did not), consistent with the idea that a pacemaker generating duration representations ran faster for auditory than for visual stimuli. The different variability of auditory and visual stimuli was attributed to differential variability in the operation of a switch of a pacemaker-accumulator clock, and experimental data suggested that such switch effects were separable from changes in pacemaker speed. Overall, the work showed how a clock model consistent with scalar timing theory, the leading account of animal timing, can address an issue derived from the classical literature on human time perception.

Data from experiments on time perception in humans have often shown that the subjective duration of a stimulus can be influenced by factors in addition to its actual physical length. For example, stimuli that are "filled" (e.g. continuous tones) are usually perceived as longer than equal-length stimuli that are "empty" (e.g. the same duration started and ended by clicks). Thomas and Brown (1974) is a modern reference for this "filled duration illusion", but Fraisse (1964) traces the result back at least to Meumann in 1896. Likewise, moving stimuli have been judged as longer in duration than static ones (Brown, 1995; Goldstone & Lhamon, 1974), presentations of familiar words were judged as lasting longer than unfamiliar ones (Witherspoon & Allan, 1985), high-pitched sounds can be judged as longer than lower-pitched ones (Cohen, Hansel, & Sylvester, 1954; but see Goldstone & Goldfarb, 1964a), and a frequent result from the classical timing literature is

---

Requests for reprints should be sent to J.H. Wearden, Department of Psychology, The University, Manchester, M13 9PL, U.K. E-mail: wearden@psy.man.ac.uk.

The experiments reported in this article were conducted in partial fulfilment of the requirements for bachelors degrees in Psychology awarded to the second, third, and fourth authors by Manchester University.

that more intense stimuli tend to be judged as longer than less intense ones (Fraisse, 1964, p. 134; Goldstone & Goldfarb, 1964a).

The present article addresses a particular subset of these effects, namely the phenomenon that "sounds are judged longer than lights" (e.g. Goldstone & Lhamon, 1974): more precisely, the result that auditory stimuli frequently appear to have longer subjective durations than do visual stimuli of the same real-time length. It is extremely important to emphasize two aspects of our work at the outset. First, we do not attempt to demonstrate that *all* auditory and visual comparisons must necessarily produce this result. Indeed, a theoretical account that we provide later suggests that exceptions probably exist, and it may even provide rules for generating them. Secondly, the aim of the present article is not merely to provide further examples of Goldstone and Lhamon's assertion, which in any case is amply supported by the material they reviewed in their article in 1974. Rather, we attempt here to employ ideas derived from a currently popular account of animal and human timing—the *scalar timing theory* of Gibbon, Church, and Meck (1984)—to address the issue of why auditory/ visual subjective length differences may occur. That is, the aim of the present article is primarily *theoretical*, using experimental data from auditory/ visual temporal comparisons (using an arbitrary pair of stimuli for which a marked auditory/ visual duration judgement is obtained) to investigate mechanisms of the hypothetical internal clock that scalar timing theory proposes.

Although scalar timing was originally developed as a model of animal timing (e.g. Gibbon, 1977) and continues to be fruitfully applied to data from animals (e.g. Church, Meck, & Gibbon, 1994; Leak & Gibbon, 1995), it has recently enjoyed some success as an account of timing in humans under a variety of procedures including interval production (Wearden, 1991a; Wearden & McShane, 1988), bisection (Allan & Gibbon, 1991; Wearden, 1991b; Wearden & Ferrara, 1995, 1996), temporal generalization (Wearden, 1992; Wearden & Towse, 1994; Wearden, Denovan, Fakhri, & Howarth, 1997a), and categorical timing (Wearden, 1995). The general principles of scalar timing as well as some of its specifics have also recently been applied to a study of age and intelligence effects in human timing (Wearden, Wearden, & Rabbitt, 1997c).

Many of these studies have involved what Wearden (1991a) called *analogue experiments*—that is, procedures used with humans that take as their methodological starting-point previous experimental work with animals. These analogue experiments have been useful in exploring the specific issue of animal–human comparisons in the domain of timing (an issue which continues to be of interest, see Wearden, Rogers, & Thomas, 1997b), but the application of scalar timing theory need not be restricted just to data obtained from methods that relate to animal psychology. More recently, studies have been conducted using techniques like verbal estimation, which have no equivalent in animals, and in general an important next step for scalar timing might be to address issues that derive from traditional work on time perception in humans. Such a step is attempted in the present article.

Many studies within the scalar timing framework with both animal and human subjects have been concerned with conformity of behaviour to the mathematical properties required by the scalar model (Gibbon et al., 1984), particularly the *scalar property* itself, the requirement that the standard deviation of underlying time representations be a constant fraction of the mean, giving rise to a constant coefficient of variation (standard deviation/ mean) as the duration timed varies. This is a form of conformity to Weber's law and was found to be almost universal in studies of timing in both animals (e.g. Church & Gibbon, 1982; Church et al., 1994), and humans (Allan & Gibbon, 1991; Wearden, 1991 a, b, 1992, 1995; Wearden & Ferrara, 1995, 1996; Wearden et al., 1997 a, b, c). More recently, however, experimental interest has also focused on the mechanics of the hypothetical internal clock (discussed later), such as how different experimental conditions might change pacemaker speed (usually just called "clock speed"), and other processes involved in clock operation.

Although details of most previous experiments on scalar timing in humans are not directly pertinent to the work here, the contents of Penton-Voak, Edwards, Percival, and Wearden (1996) do bear directly on the present study. According to scalar timing theory, the raw material for time judgements comes from a pacemaker–accumulator-type internal clock, and Penton-Voak et al. (1996) used a manipulation derived from experiments by Treisman, Faulkner, Naish, and Brogan, (1990, and replicated in Experiment 3 of the present article) in attempts to increase the speed at which the pacemaker ran. Preceding short tones (or visual stimuli) with trains of clicks increased their subjective length, in a manner consistent with the idea that the speed at which the pacemaker ran during tone or visual stimulus presentation was increased by the clicks. To anticipate results presented later, we find that the auditory stimulus we use has a greater subjective length than our visual stimulus, when the two actually have the same duration. One interpretation of this result is that the pacemaker of an internal clock of the type proposed by scalar timing theory runs faster during auditory than visual stimuli, so modality affects on timing may in fact be disguised manipulations of clock speed. However, even if this is true, clock speed differences may not be the only differences in operation of the pacemaker-accumulator clock occurring in auditory/ visual comparisons. Experiment 1 presents an initial demonstration of auditory/ visual timing differences that (a) replicates the result commonly found in the previous literature (that tones appear to last longer than visual stimuli) and (b) shows that there may be more to auditory/ visual comparisons than just differences in mean subjective length.

## EXPERIMENT 1

The technique of *temporal generalization* in humans was developed by Wearden (1991a, 1992) from an animal experiment by Church and Gibbon (1982). In the usual version of the human analogue experiment, subjects initially receive examples of a standard duration (e.g. a 400-msec tone) identified as such. Following this, they receive series of stimuli, including the standard, but also with longer and shorter stimuli, and they must simply decide whether or not each presented stimulus has the same duration as the standard (a YES or NO response), with accurate feedback being given after each trial.

The temporal generalization variant used in Experiment 1 deviates in some details from this technique and is similar to the *roving standard* method previously used in time psychophysics (Allan, 1979). The standard presented was either an auditory or a visual stimulus, and this was followed by comparison stimuli (which could differ in length from the standard, as well as possibly being in different modalities). After each comparison stimulus, the subject had to judge whether or not it was identical in duration to the standard, but no feedback was given.

The obvious focus of experimental interest was on which durations of comparison stimuli were judged as the same as the standard, particularly in the cross-modal conditions. For example, if auditory stimuli have a greater subjective length than visual stimuli of the same real duration, then *longer* visual comparison stimuli should be matched to the auditory standard and *shorter* auditory comparison stimuli matched to the visual standard. In our Experiment 1, one measure of this is the shape of the *temporal generalization gradient* (in the present case, the proportion of identifications of a stimulus as the standard plotted against difference in duration between the standard and the comparison). If auditory stimuli are subjectively longer than visual ones, this temporal generalization gradient should be skewed to the right when the standard is auditory and the comparisons visual, and to the left when the standard/ comparison modality relations are reversed. If the auditory/ visual subjective length relations were the opposite and visual stimuli appeared subjectively longer than auditory ones, then the skew of the temporal generalization gradients should, of course, be the opposite also.

## Method

### Subjects

Eleven Manchester University undergraduates participated for course credit.

### Apparatus

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

### Procedure

The auditory stimulus used was a 500-Hz tone produced by the computer speaker; the visual stimulus was a 4 × 4-cm light-blue square, centred in the middle of the monitor screen. The experiment was arranged in a series of blocks of trials, with each block consisting of the presentation of 4 examples of a standard duration, followed by 7 comparison durations. The standards and the comparison could either be auditory (aud) or visual (vis), giving rise to 4 different stimulus conditions, depending on the modality of the standard and comparisons. These were aud/ aud (i.e. auditory standard, auditory comparison), vis/ vis, aud/ vis, and vis/ aud. For illustration, consider an aud/ aud block. The standard duration for the block was chosen from a uniform distribution running from 400 to 600 msec. This value changed between blocks but was constant within a block. The auditory standard was presented 4 times, identified as such by an appropriate display on the screen, with presentations separated by intervals randomly chosen from a uniform distribution between 2000 and 3000 msec. Following this, the 7 comparison durations were presented. Their durations were the duration of the standard, whatever it was on that block, with the quantities −300, −200, −100, 0, 100, 200, and 300 (in msec) added. The seven comparison durations were randomly ordered, and the subject pressed the spacebar to produce each one in response to a prompt on the screen. This spacebar response was followed by a delay ranging from 2000 to 3000 msec and then by the comparison stimulus presentation. When the comparison stimulus terminated, subjects were asked "Did that stimulus have the same length as the standard? Press Y (YES) or N (NO) keys". No feedback was given for the response. The procedure for the other types of blocks (vis/ vis, aud/ vis, vis/ aud) was identical except for the modalities of the standard and comparison durations. Three blocks were given of each of the 4 different standard/ comparison types.

## Results

Figure 1 plots the mean proportion of YES responses (i.e. judgements that the presented comparison duration was the same as the previously presented standard) are plotted against the difference (in msec) in between the standard duration presented and the comparison stimulus judged. The upper panel shows data from the within-modality comparisons aud/ aud and vis/ vis. For both types of stimulus comparisons, the peak of YES responses was found at the 0-msec difference (i.e. when the standard and comparison really did have the same length), and the temporal generalization gradients were more or less symmetrical around this peak value. Comparisons of the average proportion of YES responses to stimuli longer than the standard with the same measure for stimuli shorter yielded no significant differences in either the aud/ aud or vis/ vis case. Inspection of the results suggests, in addition, that the vis/ vis gradient was flatter than the aud/ aud one. Analysis of variance (ANOVA) found no significant effect of condition (aud/ aud or vis/ vis), $F(1, 10) = 3.06$, but a highly significant effect of stimulus difference, $F(6, 60) = 18.41$, $p < .001$. The Modality Condition × Stimulus Difference interaction also just reached significance, $F(6, 60) = 2.25$, $p = .05$. This latter effect presumably reflects the differential flatness of the curves for the different modalities, and the fact that which modality produced more YES responses depended on stimulus difference. Another way of examining this is to compare the average number of YES responses at the 4 most extreme stimulus differences. A *t*-test showed that these differed significantly between the modality conditions, $t(10) = 4.35$, $p = .001$, confirming statistically that the vis/ vis comparisons produced more YES responses at large stimulus differences than did aud/ aud ones.

The lower panel of Figure 1 shows the cross-modal vis/ aud and aud/ vis comparisons. In these conditions, temporal generalization gradients were highly skewed around the 0-msec difference and, more importantly, skewed in different directions. For example, visual standards tended to be matched to auditory comparisons that were *shorter* than themselves (i.e. negative differences, condition vis/ aud), whereas auditory standards tended to be matched to visual comparisons that were *longer* than themselves (i.e. positive differences, condition aud/ vis). ANOVA produced significant effects of condition (i.e. aud/ vis or vis/ aud), $F(1, 10) = 14.05$, $p = .004$, stimulus difference, $F(6, 60) = 9.17$, $p < .0001$, and Condition × Stimulus Difference interaction, $F(6, 60) = 6.56$, $p < .0001$.

Other aspects of the data from cross-modal comparisons were examined by *t*-tests. Both the aud/ vis and vis/ aud gradients were significantly asymmetrical around the 0-msec difference, when the mean number of YES responses to comparisons with durations

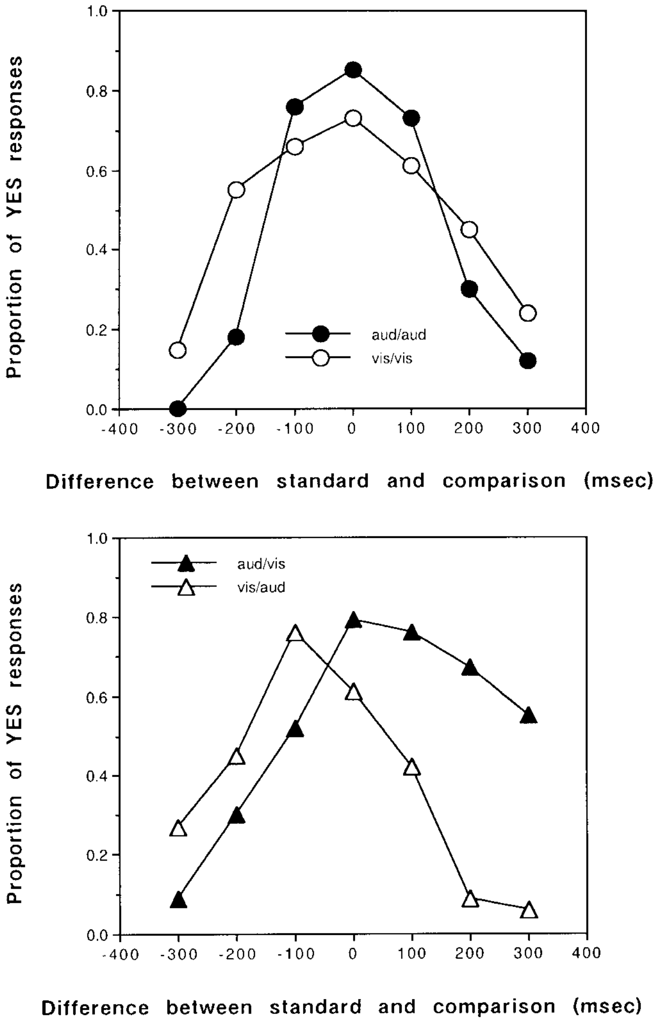

FIG. 1. Mean proportion of YES responses (identifications of a comparison stimulus as the previously presented standard duration), plotted against comparison/ standard difference from the temporal generalization study (Experiment 1). Upper panel: data from within-modality (aud/ aud and vis/ vis) comparisons. Lower panel: data from between-modality (aud/ vis and vis/ aud) comparisons.

longer and shorter than the standard were compared (aud/ vis: $t(10) = -2.81$, $p = .02$; vis/ aud: $t(10) = -3.53$, $p = .005$). They were also asymmetrical in different directions. Although the mean number of YES responses to comparisons shorter than the standard was not significant when the vis/ aud and aud/ vis conditions were compared, the number of YES responses made to comparison stimuli longer than the standard was significantly different in the different cross-modal comparisons, $t(10) = 8.09$, $p < .001$.

## Discussion

Our results suggest that the visual stimulus we used was perceived as shorter than the auditory stimulus, when durations were in fact the same. For example, a visual standard stimulus of an average duration of 500 msec was maximally matched to an auditory stimulus 100 msec shorter (lower panel of Figure 1). Conversely, an auditory stimulus was more likely to be matched to a visual stimulus longer than one that was shorter. Another effect implied by our results was that the visual stimulus representation was more variable than the auditory one. Evidence for this comes from the within-modality conditions aud/ aud and vis/ vis. The temporal generalization gradient from the vis/ vis condition was flatter than that from the aud/ aud comparison, indicating that subjects were more likely to confuse actually very different durations (e.g. those with an absolute difference of 300 msec, or an average of more than 50% of their length) when these were visual than when they were auditory.

Unlike temporal generalization gradients obtained from previous experiments from our laboratory, when standard durations were short tones (Wearden, 1992), short intervals delineated by clicks (Wearden, 1991a), or longer auditory or visual durations (Wearden et al., 1997a), the gradients obtained from the aud/ aud and vis/ vis conditions of Experiment 1 were symmetrical. Previously, durations longer than the standard have been found to be more confusable with it than durations shorter by the same amount, leading to significantly asymmetrical gradients. The reason for this difference is unclear, but one possibility is that the present Experiment 1 did not provide any feedback after the responses made to comparison stimuli (as such feedback might have allowed subjects to compensate for any auditory/ visual differences in the cross-modal conditions), whereas, in previous experiments, consistent and accurate feedback (analogous to the reinforcement in the original animal experiment by Church & Gibbon, 1982) was always given.

# EXPERIMENT 2

According to quantitative accounts of the operation of internal clocks, subjective length differences between stimuli, such as obtained in our Experiment 1, can arise in a number of ways. Suppose we consider a simple pacemaker–accumulator internal clock of the type suggested by scalar timing (e.g. Gibbon & Church, 1984), and propose that the operation of this clock differs for auditory and visual stimuli. One possibility is that the rate of the pacemaker differs for the different stimuli, running and rate, $r_a$ for auditory stimuli and $r_v$ for visual ones, with $r_a > r_v$. Obviously, for some constant stimulus length, $t$, more pulses will accumulate for auditory than for visual stimuli, giving rise to a subjective length difference in the direction found in Experiment 1. But this is not the only possibility, and another concerns the latency of operation of a hypothesized switch between the pacemaker and accumulator. According to Gibbon and Church (1984), when a stimulus to be timed begins, pulses flow from the pacemaker to the accumulator via a switched connection,but this switch has non-zero latency $l_c$ to close (i.e. allow pulses to flow) and $l_o$ to open (and cut the pacemaker–accumulator connection) when the stimulus ceases. Thus for some pacemaker rate $r$, the number of pulses accumulating in some time $t$ is $r(t - l_c + l_o)$. Subjective length of stimuli in different modalities could be different if the balance of switch latencies differed between modalities (e.g. the switch closed faster, or opened more slowly, with auditory rather than visual 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 across auditory and visual modalities, 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 upon the difference $l_o - l_c$ being greater than zero. Even if the absolute values of switch opening and closing were different for auditory and visual modalities (if, for example, both the onset and offset of stimuli to be timed were registered more rapidly for auditory rather than for visual stimuli), an intercept difference between modalities would only be observed if the *difference* between opening and closing latencies varied.

The contrary case is the one in which pacemaker rate is the same for auditory and visual stimuli, in which case only the second additive component of the function given above plays a role, and intercept, but not slope, differences between modalities would be observed.

In Experiment 2, subjects were required to estimate verbally the lengths of 10 stimuli presented in either auditory or visual modalities. Their judgements could be arranged as plots of mean estimate versus real stimulus length, enabling effects of modality on slope and intercept of the estimate function to be measured. As Penton-Voak et al. (1996) note, although in principle slope and intercept differences are straightforward to observe, an important practical requirement is the need for a substantial range of stimulus lengths, thus producing a sufficient amount of independent variable variance for statistical discrimination with practicable numbers of observations. Their Experiment 3b suggested a set of time values that could discriminate between the two potential effects of the clicks they used to “speed up the internal clock”, and we used the same time values in Experiment 2.

## Method

### Subjects

Fourteen Manchester University undergraduates took part.

### Apparatus

Apparatus and other details of setting were as Experiment 1.

### Procedure

The auditory and visual stimuli were the same as in Experiment 1. Each trial involved the presentation of a single stimulus, either in the auditory or visual modality. The stimulus was presented after a spacebar press in response to an on-screen prompt and occurred between 2000 and

3000 msec after the press. Stimulus lengths were 77, 203, 348, 461, 582, 767, 834, 958, 1065, and 1183 msec, and these values were presented either as auditory or as visual stimuli, generating a block of 20 trials (modality × stimulus length). Within the block, the 20 stimuli were randomly arranged and presented successively until all 20 had been used. Another random order was arranged for the next block, and so on. Five blocks were given in a single experimental session. When each stimulus had been presented, subjects were required to type in an estimate of its length in msec. They were prompted to do this by an on-screen display, which also reminded them that 1000 msec = 1 sec. No feedback was given after the response. Subjects were told that all durations were between 50 and 1500 msec in length.

## Results

Verbal estimates produced by individual subjects were first filtered to discard all values outside the range specified in the instructions (50 to 1500 msec). This resulted in the loss of just a few percent of data and was done to eliminate errors due to mistypings (e.g. "10,000" for "1000").

Mean verbal estimates produced for the visual and auditory stimuli are shown, plotted against real stimulus length, in the upper panel of Figure 2. Inspection of the data suggested that although estimates of the lengths of both types of stimuli increased as an orderly, and approximately linear, function of real duration, the auditory stimuli were judged as longer. ANOVA supported these suggestions, showing a highly significant effect of stimulus modality, $F(1, 13) = 62.54$, $p < .0001$, and stimulus duration, $F(9, 117) = 162.95$, $p < .001$. Furthermore, a significant Modality × Duration interaction, $F(9, 117) = 5.79$, $p < .001$, confirmed the impression obvious on inspection of the data that the modality effects were more marked at the longer stimulus durations than at the shorter ones.

As mentioned above, differences in subjective duration between modalities could be due either to pacemaker-speed differences that would affect slopes (and intercepts in some cases), or to switch latency differences, which affect only intercepts. One way of testing these two possibilities directly is to perform a linear regression on the verbal estimates produced by individual subjects in the two modality conditions, then use the individual slopes and intercepts for between-modality comparisons. When this was done, it was found that 13 out of the 14 subjects produced a steeper slope with auditory judgements than with visual ones. The mean slope for the auditory stimuli was 0.95, mean for visual stimuli was 0.77, and this difference was significant by a *t*-test, $t(13) = 5.19$, $p < .001$, but differences in intercepts between modalities were not significant (mean auditory = 56.9 msec, mean visual = 30.65 msec; $t(13) = 1.31$).

The lower panel of Figure 2 shows the coefficients of variation (standard deviation/mean) of verbal estimates from the different conditions. Inspection suggests that, overall, coefficients of variation tended to decline with increases in stimulus length (except for low values at the shortest stimulus length), and that coefficients of variation were higher for visual than for auditory stimuli—that is, the estimates for visual stimuli were more variable around their mean. Both these suggestions were confirmed statistically, the first by a significant effect of stimulus length, $F(9, 117) = 8.3$, $p < .001$, the second by a significant modality effect, $F(1, 13) = 8.06$, $p = .014$. However, there was no significant Modality × Stimulus Length interaction for coefficients of variation, $F(9, 117) = 1.23$.

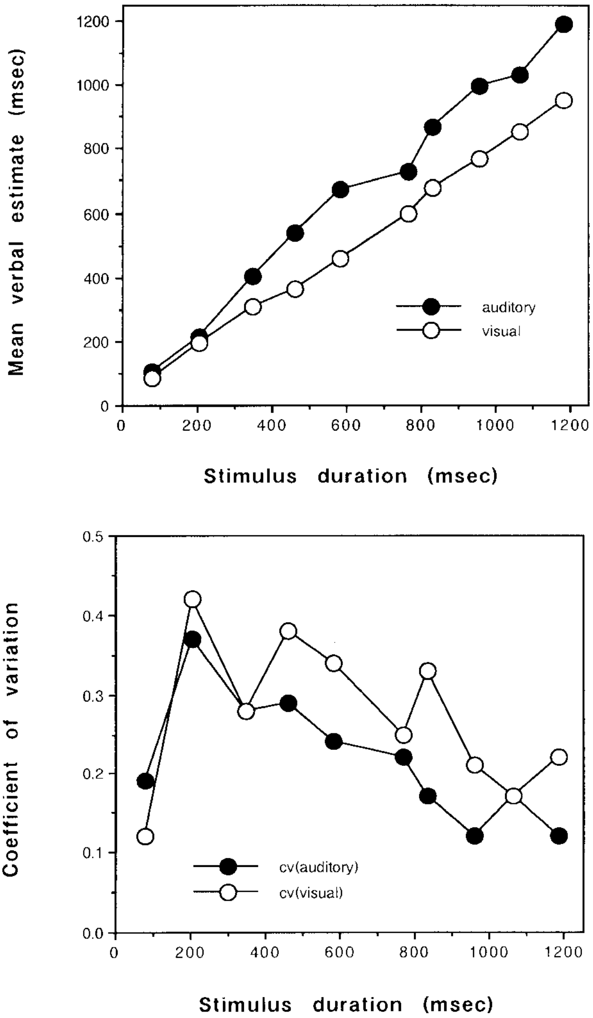

FIG. 2. Upper panel: mean verbal estimates of duration for auditory (filled circles) and visual (open circles) from Experiment 2, plotted against stimulus length. Lower panel: coefficients of variation (standard deviation/ mean) of estimates for auditory and visual stimuli from Experiment 2, plotted against stimulus length.

## Discussion

The results of Experiment 2 confirmed and extended those obtained, using a different method, in Experiment 1. In Experiment 2, the auditory stimulus was judged as longer (upper panel of Figure 2) and less variable (lower panel of Figure 2) than the visual stimulus of the same real length. Furthermore, comparison of slopes and intercepts derived from regression of individual-subject data showed that the modality differences affected the *slope* of the function relating mean verbal estimate to stimulus length, rather than intercepts. This effect is exactly the one predicted using the idea that the pacemaker of a pacemaker–accumulator internal clock ran faster with the auditory stimulus than with the visual one, and that switch latency differences between modalities were zero or negligibly small.

Our results have some parallels in the previous literature. For example, an increasing difference between duration judgements with auditory and visual modalities as stimulus length increases (e.g. our Figure 2, upper panel) has been reported by Goldstone and Goldfarb (1964a; for a particularly clear example, see Figure 1, p. 373), Stevens and Greenbaum (1966), and Walker and Scott (1981). These results are, obviously, consistent with our hypothesis of different pacemaker speeds for auditory and visual stimuli.

If the difference in mean subjective duration usually found between auditory and visual stimuli can be attributed to pacemaker speed differences, where does the difference in relative variability come from? Within the framework of the pacemaker–accumulator clock model there are two distinct possibilities. The first is that differences in pacemaker speed are *themselves* the cause of variability differences. The pacemaker of the internal clock proposed by scalar timing is usually assumed to behave like a Poisson emitter—that is, a process that emits pulses at random but at some constant rate on average. Gibbon (1977) discusses the mathematics of such Poisson timers and shows that *slower* pacemaker rates will, by themselves, lead to more variable temporal estimates—so, for example, in the present case we would expect visual stimuli to have more variable representations in part because of the putative slower pacemaker speed with which they are associated. Within the framework of scalar timing, however, it has usually been argued (e.g. Gibbon & Church, 1984) that differences in pacemaker speed are not the principle sources of variability in the system, although they may make some small contribution to variability.

This leads to the second possibility, that the differential variability of auditory and visual stimulation in our experiments does not arise from pacemaker speed differences per se, but from some other, independent source. Figure 3 sketches a model that proposes exactly this.

Figure 3 shows the proposed operation of the pacemaker–accumulator clock for auditory and visual stimuli. The different stimulus types are hypothesized to produce two effects: first, the pacemaker runs faster for auditory than for visual stimuli; second, the latency of operation of the switch connecting the pacemaker and accumulator is proposed to be more variable in the visual than auditory case. Thus the onset and offset of visual stimuli are registered more variably than onset and offset of auditory stimuli, thus rendering the number of pulses accruing during stimulus presentation more variable from one trial to another (as well as being different in overall mean as a result of slower pacemaker speed). Appealing to switch processes in the scalar timing model, as well as possible changes in pacemaker speed, to account for data obtained in timing experiments with humans has precedents in some previous work (e.g. Allan, 1992), although no previous article has proposed our particular combination of effects as an account of cross-modal phenomena.

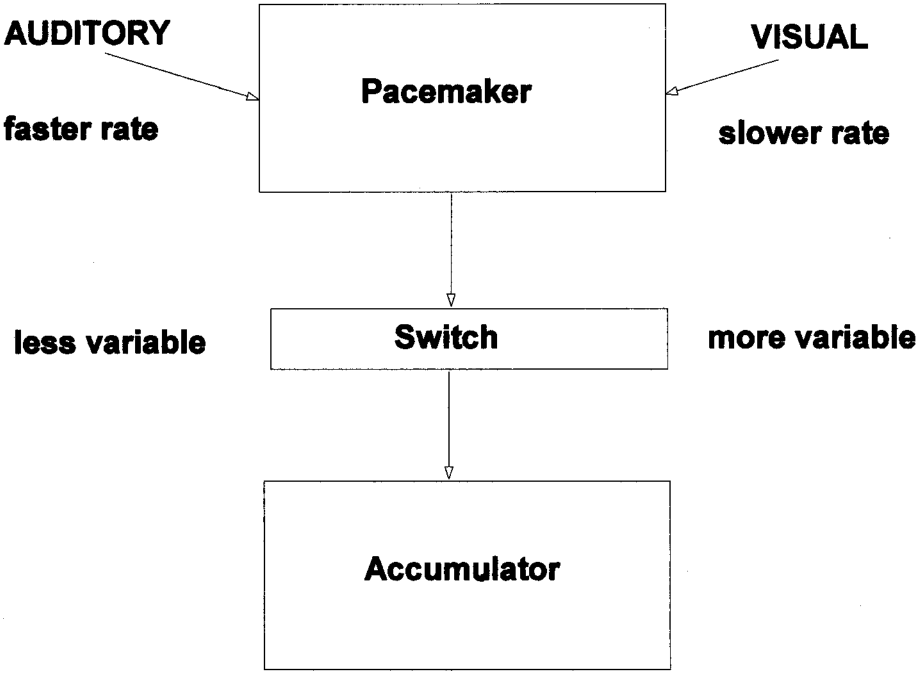

FIG. 3. Model of internal clock operation for auditory and visual stimuli. See text for details.

## EXPERIMENT 3

How can the model in Figure 3 be evaluated? One of its predictions, in contrast to the pacemaker speed hypothesis of variability differences, is that it should be possible to manipulate pacemaker speed and variability separately. The first hypothesis advanced above, however, suggests that these must necessarily vary together. Penton-Voak et al. (1996) showed that preceding auditory and visual stimuli by trains of clicks increased their subjective duration in a manner consistent with the idea that the pacemaker had been speeded up by the clicks (i.e. a slope difference between click and no-click conditions—see also Treisman et al., 1990, for a similar result). The present Experiment 3, below, repeated this manipulation with within-subject comparisons. In Experiment 3, auditory or visual stimuli were presented either alone or preceded by a 5-sec train of 5-Hz clicks, and verbal estimation of length was required. The two no-click conditions constitute a smaller-scale replication of the present Experiment 2, but of additional interest is the possible effect of clicks.

The focus of interest is not only on the possibility that clicks will manipulate the subjective length of both the auditory and visual stimuli (thus replicating Penton-Voak et al., 1996), but also on two more specific questions. First, is the effect of the clicks similar for both auditory and visual stimuli, thus possibly implicating a common timer for both? Secondly, if mean subjective length is manipulated, how is estimate variability affected? If the rate of the pacemaker itself produces the variability, then changing mean subjective length should affect variability. On the other hand if, as the model in Figure 3 suggests, the differential variability of auditory and visual stimuli comes from a source different from the subjective length difference (i.e. a switch rather than pacemaker effect), then variability differences between auditory and visual stimuli should persist even when mean subjective length is altered.

## Method

### Subjects

Sixteen Manchester University undergraduates took part.

### Apparatus

As Experiments 1 and 2.

### Procedure

Subjects experienced a single experimental session, which involved stimulus presentations of 4 different types. Six different durations (77, 348, 582, 767, 958, and 1183 msec) were presented either as the 500-Hz tone (auditory) or the blue square (visual) used in Experiments 1 and 2. The stimuli were preceded either by 5 sec of silence (no click, NC, conditions) or by 5 sec of auditory clicks (click, C, conditions), where the auditory clicks were 1000-Hz tones, 10 msec in duration, produced by the computer speaker, with an onset-to-onset interval resulting in a 5-Hz click frequency. The combination of C/ NC conditions, two modalities, and the six durations resulted in a block of 24 stimuli, and within the block these were arranged in a random order and successively presented until all 24 had been used. 4 blocks of stimuli constituted the session. The 4 types of stimulus events used within a block were designated as VIS/ NC (visual stimulus preceded by silence), VIS/ C (visual preceded by clicks), AUD/ NC, and AUD/ C. After each stimulus had been presented, subjects were required to type in an estimate of its duration in msec, but no feedback was given after this response. All experimental details not mentioned explicitly were the same as for Experiment 2.

## Results

As in Experiment 2, data were filtered to remove the few percent of estimates outside the specified range (50 to 1500 msec). We initially subjected the mean verbal estimates to a three-way ANOVA, with stimulus duration, stimulus modality, and presence or absence of clicks as the three factors. All three main effects were significant (stimulus duration, $F(5, 75) = 226.5$, $p < .001$; modality, $F(1, 15) = 30.12$, $p < .001$; clicks, $F(1, 15) = 44.21$, $p < .001$). There was a significant Modality × Stimulus Duration interaction, $F(5, 75) = 8.23$, $p < .001$, and a Clicks × Stimulus Duration interaction, $F(5, 75) = 8.24$, $p < .001$, which suggest that modality effects and click effects on mean estimates were reflected in slope differences (a conclusion supported by analysis presented later), whereas the Clicks × Modality interaction was not significant, $F(1, 15) = .57$, $p = .50$, suggesting that the effect of clicks was the same for both auditory and visual stimuli. The three-way interaction was also non-significant, $F(5, 75) = .40$, $p = .85$.

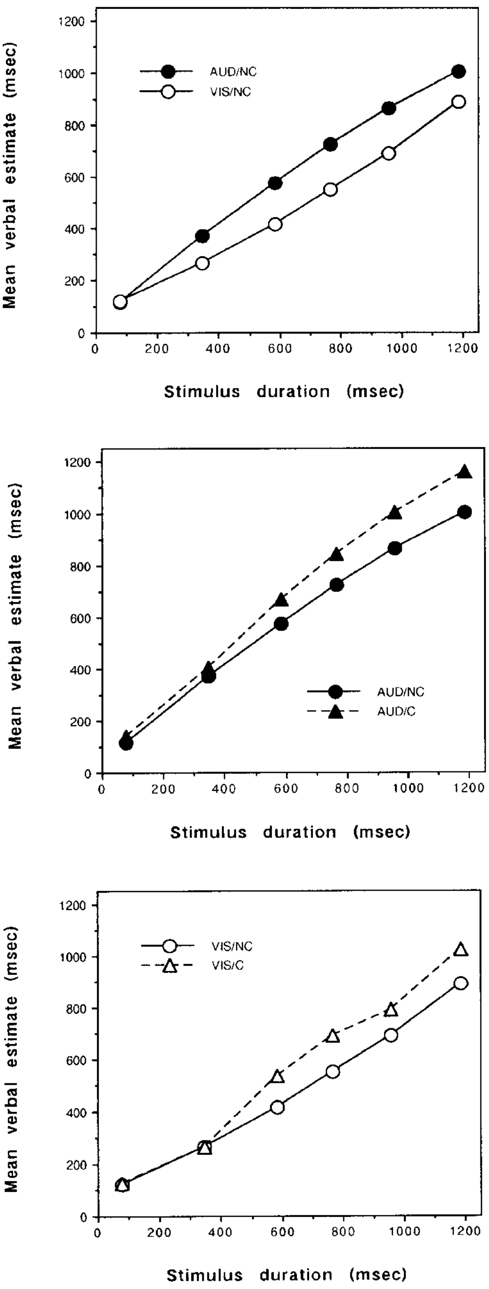

FIG. 4. Mean verbal estimates plotted against stimulus duration from Experiment 3. Conditions are visual/ no clicks (VIS/ NC), auditory/ no clicks (AUD/ NC), and visual and auditory stimuli preceded by clicks. Upper panel: judgements of auditory and visual stimulus durations without clicks. Centre panel: effect of clicks on estimates of auditory stimulus length. Bottom panel: effect of clicks on estimates of visual stimulus length.

Although the three-way ANOVA produced a number of interpretable effects, exposition of the results is probably clearer when two-way ANOVAs are used to address different specific questions. One of these is whether the between-modality mean estimate difference found in Experiment 2 was replicated. The upper panel of Figure 4 shows the mean estimates from the VIS/ NC and AUD/ NC (i.e. when both types of stimuli were preceded by silence)—the relevant conditions for this comparison. Obviously, overall the auditory stimuli produced longer mean estimates than the visual ones. ANOVA found significant effect of modality, $F(1, 15) = 20.07$, $p < .0001$, and of stimulus duration, $F(5, 75) = 201.29$, $p < .0001$, and a Modality × Stimulus Duration interaction, $F(5, 75) = 5.92$, $p < .001$, with the latter effect presumably reflecting the increased difference between the estimates for stimuli in different modalities at the longer stimulus durations compared with the shorter ones. An analogous effect was found when comparing the two stimulus types when both were preceded by clicks (VIS/ C and AUD/ C), although these are not shown in Figure 4. This comparison yielded significant effects of modality, $F(1, 15) = 36.41$, $p < .0001$, and of stimulus duration, $F(5, 75) = 179.65$, $p < .0001$, and a Modality × Stimulus Duration interaction, $F(5, 75) = 4.15$, $p < .01$. Thus, overall, auditory stimuli were judged longer than visual ones, whether both were preceded by silence or both by clicks, and differences tended to be greater at longer stimulus durations.

The other type of comparison of interest is between click and no-click conditions. These are shown in Figure 4 separately for auditory stimuli (centre panel), and visual stimuli (bottom panel). Obviously, in both cases, preceding the stimuli by a 5-sec train of 5-Hz clicks increased their subjective length, and effects were larger at longer stimulus durations. These observations were confirmed by ANOVA of data from the AUD/ NC and AUD/ C conditions, which found a significant effect of clicks, $F(1, 15) = 28.35$, $p < .0001$, and of stimulus duration, $F(5, 75) = 190.01$, $p < .0001$, and a significant Click × Stimulus Duration interaction, $F(5, 75) = 3.81$, $p = .004$. Comparisons of the VIS/ C and VIS/ NC conditions produced analogous effects, with a significant effect of clicks, $F(1, 15) = 37.46$, $p < .0001$, and of stimulus duration, $F(5, 75) = 166.56$, $p < .0001$, and an interaction, $F(5, 75) = 4.96$, $p = .001$. Thus the clicks increased the subjective length of both auditory and visual stimuli and increased length in a similar way for both.

To test whether the *magnitude* of the click effect differed between modalities, we constructed, for each subject, the *difference* between a particular stimulus duration estimate with and without clicks for both modalities and then compared these differences. There was now no significant modality effect, $F(1, 15) = .47$, nor any significant Modality × Stimulus Duration interaction, $F(5, 75) = .40$, but there was a significant effect of stimulus duration, $F(5, 75) = 8.24$, $p < .0001$. Thus, according to this analysis, the clicks increased verbal estimates equally for both modalities but increased them more at longer stimulus durations.

To test whether the click/ no click or modality differences were manifested as slope or intercept differences, verbal estimates from individual subjects in all 4 conditions were regressed against stimulus duration, and the resulting slope and intercept values compared by $t$-tests. Cross-modal comparisons showed that auditory stimuli produced higher slopes than visual ones both when stimuli were preceded by silence (mean auditory slope = .81, mean visual slope = .69; $t(15) = 3.21$, $p < .01$) and by clicks (mean auditory slope =

.95, mean visual slope = .82; $t(15) = 3.13$, $p < .01$). Intercept differences from the cross-modal comparison just reached significance when stimuli were not preceded by clicks (mean auditory intercept = 81.5 msec, mean visual intercept = 37.3 msec; $t(15) = 2.16$; $p = .05$), but they were significant when stimuli were preceded by clicks (mean auditory intercept = 91.13 msec, mean visual intercept = 34.29 msec; $t(15) = 2.6$, $p = .02$). However, the effects of clicks versus no clicks were pure slope effects, with clicks increasing slope when preceding either visual ($t(15) = 3.74$, $p < .01$) or auditory, $t(15) = 5.87$, $p < .001$, stimuli, but having no significant effects on intercepts (auditory: $t(15) = .57$; visual: $t(15) = -.18$). Overall, therefore, both cross-modal and click manipulations produced much more reliable slope than intercept effects.

Coefficients of variation of estimates (standard deviation/ mean) are shown in Figure 5. The first question of interest is whether the auditory/ visual variability difference noted in Experiments 1 and 2 was replicated in both the no-click and the click conditions. Consider first comparisons of auditory and visual stimuli without clicks, shown in the upper left panel. Inspection of the data suggests that the visual stimuli produced relatively more variable estimates than did auditory ones, and this was confirmed by a significant modality effect, $F(1, 15) = 10.4$, $p = .006$, but neither stimulus duration nor Duration × Modality interaction were significant. The cross-modal comparison when both stimuli were preceded by clicks (i.e. the AUD/ C and VIS/ C conditions) likewise produced a significant modality effect, $F(1, 15) = 32.2$, $p < .0001$, and a significant effect of stimulus duration, $F(5, 75) = 5.12$, $p < .001$, but the interaction was not significant (left lower panel of Figure 5).

The second question is whether the clicks, which significantly changed mean subjective length of the stimuli, also affected variability. The two other panels of Figure 5 show coefficients of variation from click and no click comparisons for auditory (top right panel) and visual (bottom right panel) stimuli. For visual stimuli there was no effect of clicks on coefficient of variation of visual stimulus estimates, $F(1, 15) = 1.28$, but there was a significant effect of stimulus duration, $F(5, 75) = 2.92$, $p = .02$. The Modality × Stimulus Duration interaction was not significant, $F(5, 75) = .80$. The analysis of click effects on auditory stimuli produced an identical pattern of results, with no effect of clicks, $F(1, 15) = 2.63$, a significant effect of stimulus duration, $F(5, 75) = 4.1$, $p = .002$, and no significant Click × Duration interaction, $F(5, 75) = 1.40$.

Overall, therefore, visual stimuli produced relatively more variable estimates than did auditory ones, whether or not they were preceded by clicks, but the presence of clicks by itself did not change variability significantly. In addition, coefficients of variation of verbal estimates generally declined significantly, albeit slightly, when stimulus duration increased.

## Discussion

Results from Experiment 3 replicated those from Experiment 2 in all important details. Estimates were longer and relatively less variable when auditory stimuli were used compared to visual ones, and the auditory/ visual difference produced significant slope differences between the functions relating mean estimate to real durations, suggesting a pacemaker–speed multiplicative effect. In Experiment 3 there was also a weaker cross-

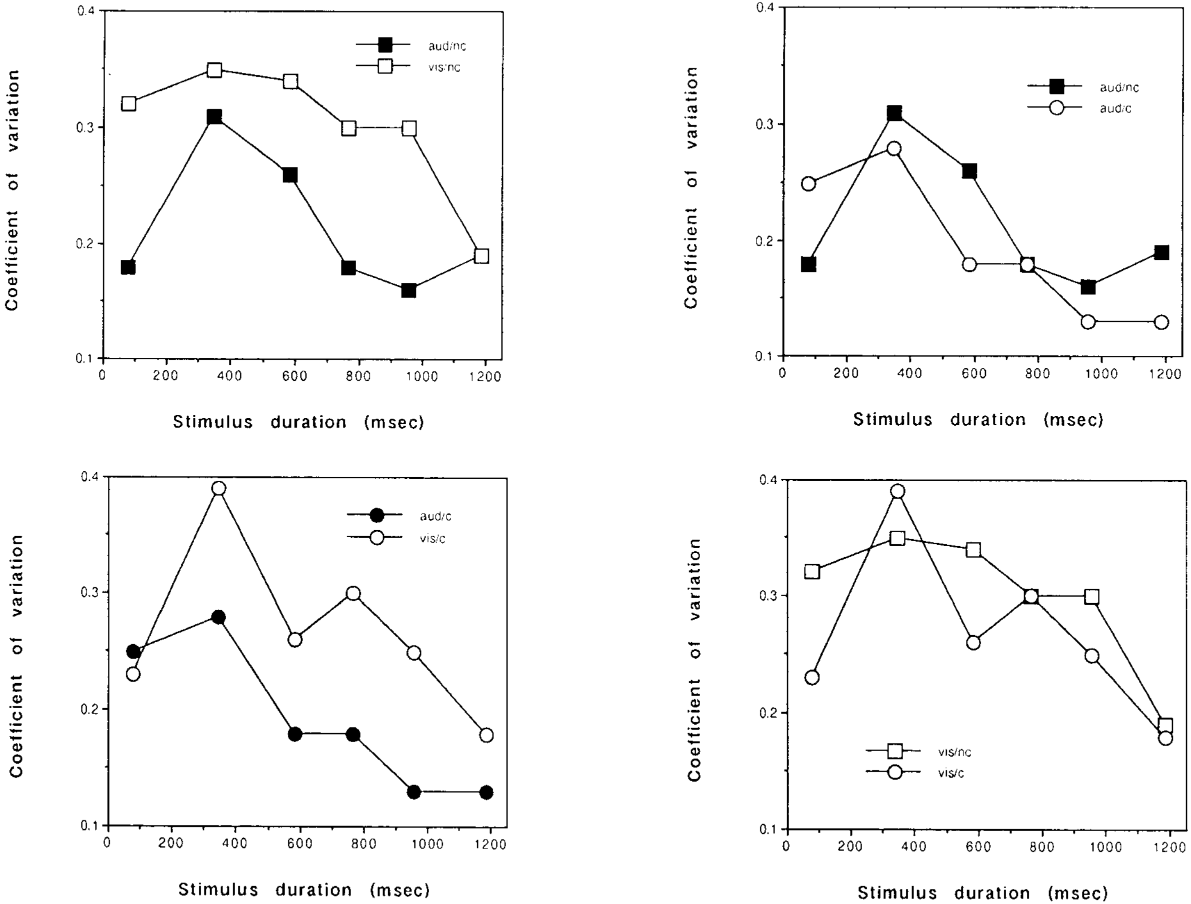

FIG. 5. Coefficients of variation from the different conditions of Experiment 3, plotted against stimulus length. Left column shows cross-modal comparisons (without clicks: upper panel; with clicks: lower panel). Right column shows effect of clicks (auditory stimuli: upper panel; visual stimuli: lower panel).

modal effect on intercept, which just reached significance when stimuli were not preceded by clicks but was clearly significant when they were. According to the equation for pulse accumulation presented above, this suggests a small effect of a difference in switch opening and closing latencies between the auditory and visual condition in Experiment 3, although this difference was not significant in Experiment 2. Furthermore, the equation suggests that any switch latency differences between auditory and visual stimuli would be more marked under conditions with higher pacemaker speed, as pacemaker speed multiplies the difference, thus predicting its clearer manifestation when stimuli were preceded by clicks in Experiment 3, which appear to increase pacemaker speed. The additional contribution of Experiment 3 was to show that click trains increased subjective durations of both visual and auditory stimuli, with the type of effect (a slope change) and the magnitude of the effect being the same for both stimulus modalities. The click effects reported here thus provide a within-subject replication of the ‘‘speeding up the clock’’ results in the various conditions of Experiment 3 in Penton-Voak et al. (1996).

The effects of clicks and modality on variability of verbal estimates supported the proposition of the model outlined in Figure 3 that variability and mean verbal estimate effects were due to separate mechanisms. While both presence and absence of clicks and a change between auditory and visual stimuli produced changes in mean verbal estimates, only the modality difference produced significantly different coefficients of variation (left panels of Figure 5). This suggests that the clicks did not significantly affect switch variability for a stimulus of a particular type (auditory or visual), although the clicks did produce a slope effect on mean verbal estimates, as predicted by their causing a change in pacemaker speed.

## GENERAL DISCUSSION

As mentioned in the introduction to this article, numerous previous studies have found the same type of auditory/ visual difference in mean subjective time that we find in our experiments above: in the overwhelming majority of studies, ‘‘sounds are judged longer than lights’’ (Goldstone & Lhamon, 1974). This conclusion applies across a variety of procedures and stimulus durations (as in Goldstone & Lhamon, 1972). For example, it is found when people judge whether or not a single stimulus was 1 sec long (Goldstone, Boardman, & Lhamon, 1959), when two stimuli (auditory and visual) are compared (Goldstone, & Goldfarb, 1964b), when people produce or reproduce the lengths of tones and lights (Goldstone, 1968; Walker & Scott, 1981), when magnitude estimation methods are used (Stevens & Greenbaum, 1966), and in a recent study using bisection methods by Penney, Meck, Allan, and Gibbon (in press)—work that will be discussed more fully later. A rare exception to the usual auditory/ visual effect was found by Brown and Hitchcock (1965) in a reproduction study where durations ranging from 1 to 17 sec were used. However, as this study used stimulus durations that were much longer than those in most other work, it is possible that subjects simply based their reproductions on values obtained by chronometric counting of the appropriate stimulus length, so this result may not alter the general conclusion that can be drawn from other studies, including our own, of auditory/ visual time differences.

It is much less clear whether previous studies have obtained the difference in variability between auditory and visual stimulus duration judgements that is such a striking feature of our data, but the problem may simply arise from the fact that few other studies have provided data on judgement variability, as opposed to measures of mean performance. However, two studies that do provide data on variability generally support our findings. Goldstone et al. (1959) reported a measure that is probably the inter-quartile range of judgements (their value Q in Tables 1 to 3 of their article), and this measure was consistently higher for visual than for auditory stimuli. Magnitude estimation and production data also usually yielded higher standard deviations for lights than for tones in Stevens and Greenbaum's (1966) study (e.g. their Table 2, p. 444; but see their Table 1, p. 443). On the other hand, in Walker and Scott's (1981) work, which involved making a motor response of the same subjective durations as a light, a tone, or a compound of the two, results on variability were mixed.

The model outlined in Figure 3 might be elaborated to produce conditions in which the usual auditory/ visual difference in mean duration judgement was reduced or reversed. If we follow Penton-Voak et al. (1996) and assume that pacemaker speed varies with arousal, then highly arousing visual stimuli might be expected to result in higher pacemaker speed (and thus longer duration judgements) than mildly arousing auditory ones. Stimulus intensity itself may manipulate arousal, and, indeed, some previous studies have found that intensity manipulations can alter the subjective length of a stimulus. For example, Goldstone and Goldfarb (1964a) found that brighter lights were judged as longer than dimmer ones, and Walker and Scott (1981) found that reducing auditory intensity reduced the subjective length of the stimulus. It seems likely, therefore, that some auditory/ visual stimulus pair for which the usual subjective length difference was reduced or reversed could be found by manipulating relative stimulus intensity.

Our model and the data supporting it from Experiment 3 suggest that pacemaker speed and switch variability effects are separable, so manipulation of stimulus characteristics might produce auditory stimuli that are as variable, or more variable, in perceived duration as visual ones. One such potentially important factor is the energy distribution at stimulus onset and offset, and it is possible that an auditory stimulus that was "ramped" to increase gradually in volume at onset and decrease gradually at offset might produce time judgements that are more variable than those made of the kind of auditory stimuli we used—perhaps even as variable as judgements of visual stimuli. However, such ramped stimuli entail the problem of deciding exactly what their real time length is—for example, are onsets and offsets measured in terms of stimulus onset and offset, whether detectable by the subject or not, or when the stimulus energy exceeds then later falls below the subject's detection threshold?

Overall, therefore, although the stimuli we used in our experiments were arbitrary, the framework shown in Figure 3 may have some generality as a testable model of future experiments on timing of auditory and visual stimuli, as well as comparisons of duration judgements made to stimuli in other modalities.

Given that judgements of the visual and auditory stimuli we used differed, as they did in the other studies reviewed above, what are the implications of our data for the question of whether there are separate pacemaker-accumulator systems for auditory and visual stimuli or a single common one? Some features of our results show that if there are two separate timing systems, they must be similar in a number of respects. Experiments 2 and 3 showed that verbal estimation of the duration of both auditory and visual stimuli produced output that translated into an approximately linear relation between mean estimate and real time (see Figures 2 and 4), suggesting that both putative clocks represent time in terms of linear accumulation of pulses (albeit probably at different absolute rates). Another manipulation that had a similar effect was preceding both auditory and visual stimuli by a 5-sec click train. In the case of both auditory and visual stimuli, this manipulation increased mean verbal estimates, and the degree and type of increase was not significantly different for the two types of stimuli. If these results and those of other work (Penton-Voak et al., 1996; Treisman et al., 1990) are compared, it appears that both the putative auditory and visual pacemakers increase their rates in a similar way in response to clicks. In view of these similarities, it may be more theoretically parsimonious to assume a common pacemaker–accumulator system for both types of stimuli, and one that also underlies timing of other events, such as response latencies (Penton-Voak et al., 1996), which are also affected by click trains in an analogous way to stimulus duration estimates.

Studies of the simultaneous timing of auditory and visual stimuli might be thought to be critical in deciding whether there are two pacemaker–accumulator systems or one, as the same single pacemaker cannot run at two different rates at the same time. Walker and Scott (1981) conducted a study (their Experiment 1) where subjects held down a button for the same length of time as the duration of a previously presented stimulus, which was either auditory, visual, or an auditory-visual compound. The visual stimulus was reproduced as shorter than the auditory one, but the auditory–visual compound was judged to have the same length as the auditory stimulus alone. Their Experiment 4 also included a condition in which subjects made separate responses for the auditory and visual stimuli, even in cases where they were presented in compound. In the compound condition, the visual part of the compound was judged to have the same length as the auditory part (although a visual stimulus alone was, as in their Experiment 1, judged shorter). Taken overall, Walker and Scott's study tends to support the idea of a single pacemaker, the rate of which is subject to "auditory dominance" (i.e. when an auditory–visual compound is presented, pacemaker rate is determined by the auditory stimulus).

A recent article by Penney et al. (in press) also discusses judgements of auditory and visual stimulus duration by humans in the context of scalar timing theory. Their article reports data collected with a bisection method (somewhat similar to that used in Allan & Gibbon, 1991, and in Wearden & Ferrara, 1995, 1996), where student subjects were initially presented with examples of "Short" and "Long" standards (e.g. 3 and 6 sec, or 4 and 12 sec), then classified a range of durations (the standards and some other durations in-between) in terms of whether each stimulus had the same duration as either the Short or the Long standard. In Penney et al.'s Experiment 1, different subject groups received *either* auditory *or* visual stimuli, but not both, and bisection performance did not differ significantly between the two groups. In Experiment 2, however, subjects experienced *both* stimulus modalities in the same session, and this time a distinct auditory/visual difference, in the direction of auditory stimuli being subjectively longer than visual ones, was found. In Penney et al.'s study, however, the bisection task was never *explicitly* cross-modal (i.e. subjects were not required to make judgements relating an auditory comparison stimulus to a visual standard, or vice versa), unlike the temporal generalization experiment reported in the present article. Penney et al. (in press) hypothesized that when people received a single stimulus modality, any subjective length difference between auditory and visual stimuli would be undetectable, as all the stimuli (i.e. all auditory or all visual) would have been judged using the same pacemaker speed. When the two modalities are presented together, on the other hand, a process of "memory mixing" is proposed. For example, suppose that the subjective lengths of the standard "Short" auditory and visual are mixed together to form a common "Short" standard, and the "Long" auditory and visual standards are likewise mixed to form a common "Long". Now, auditory stimuli will tend to be subjectively long relative to this common standard, and visual stimuli subjectively short, so auditory/ visual differences in bisection can be observed.

In theoretical terms, Penney et al. (in press) accounted for auditory–visual subjective length differences in terms of a multiplicative effect resulting from different pacemaker speeds and proposed that the visual pacemaker speed was 92% of the auditory speed. Furthermore, in their mathematical model of bisection performance they used different coefficients of variation of remembered values of the standard durations, with parameter values such that the representation of the visual standards was more variable than that of the auditory standards. The treatment of our data, which derive from experiments that are in some ways logically simpler than those of Penney et al. (in press), led to very similar conclusions, but our treatment was guided by the results of experimental manipulations (e.g. slope versus intercept effects in Experiments 2 and 3, the ability to manipulate pacemaker speed separately from variability in Experiment 3) rather than by curve fitting, as in Penney et al. (in press). In spite of a multitude of procedural differences (in method used and durations timed, as well as different physical stimuli for the auditory and visual signals) between our study and that of Penney and colleagues, the two articles together suggest a consistent answer to the question of why "tones are judged longer than lights": the hypothesized internal clock runs faster for auditory than for visual stimuli, and this "clock speed" difference is the main source of differences in subjective duration. Our Experiment 3 suggested, in addition, a small contribution to differences in subjective duration of differences in the balance of switch closure and opening latencies for the auditory and visual stimuli.

In conclusion, not only do our experiments present a consistent body of data on auditory/ visual stimulus timing judgements, showing reliable effects of modality on both mean and variability of subjective duration, but we also link the data obtained (and by implication previous data also) to some contemporary ideas about the operation of a hypothesized internal clock in humans, derived from work by Gibbon et al. (1984). The internal clock is only one part of the scalar timing system, but the apparently separable effects on clock speed and switch variance are sufficient to account for the patterns obtained in our data. Scalar timing theory, initially developed as an account of animal timing, can therefore not only be extended to human timing in experimental procedures modelled on animal experiments (e.g. Wearden, 1992, 1995), it can also contribute to understanding of results in the classical literature on human timing.

# REFERENCES

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: Kluwer.

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 and Psychophysics*, *57*, 105–116.

Brown, D.R., & Hitchcock, L. (1965). Time estimation: Dependence and independence of modality-specific effects. *Perceptual and Motor Skills*, *21*, 727–734.

Church, R.M., & Gibbon, J. (1982). Temporal generalization. *Journal of Experimental Psychology: Animal Behavior Processes*, *8*, 165–186.

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.

Cohen, J., Hansel, C.E.M., & Sylvester, J.D. (1954). Interdependence of temporal and auditory judgments. *Nature*, *174*, 642–646.

Fraisse, P. (1964). *The psychology of time*. London: Eyre and Spottiswoode.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, *84*, 279–325.

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: Lawrence Erlbaum Associates, Inc.

Gibbon, J., Church, R.M., & Meck, W. (1984). Scalar timing in memory. In J. Gibbon and L. Allan (Eds.), *Annals of the New York Academy of Sciences, Vol. 423: Timing and time perception* (pp. 52–77). New York: New York Academy of Sciences.

Goldstone, S. (1968). Production and reproduction of duration: Intersensory comparisons. *Perceptual and Motor Skills*, *26*, 755–760.

Goldstone, S., Boardman, W.A., & Lhamon, W.T. (1959). Intersensory comparison of temporal judgments. *Journal of Experimental Psychology*, *57*, 243–248.

Goldstone, S., & Goldfarb, J.L. (1964a). Auditory and visual time judgment. *Journal of General Psychology*, *70*, 369–387.

Goldstone, S., & Goldfarb, J.L. (1964b). Direct comparison of auditory and visual durations. *Journal of Experimental Psychology*, *67*, 483–485.

Goldstone, S., & Lhamon, W.T. (1972). Auditory-visual differences in human temporal judgment. *Perceptual and Motor Skills*, *34*, 623–633.

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.

Leak, T.M., & Gibbon, J. (1995). Simultaneous timing of multiple intervals: Implications of the scalar property. *Journal of Experimental Psychology: Animal Behavior Processes*, *21*, 3–19.

Penney, T., Meck, W.H., Allan, L.G., & Gibbon, J. (in press). Human temporal bisection: Mixed memories for auditory and visual signal durations. In C.E. Collyer & D.A. Rosenbaum (Eds.), *Timing of behavior: Neural, computational, and psychological perspectives*. Cambridge, MA.: MIT Press.

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.

Stevens, S.S., & Greenbaum, H.B. (1966). Regression effect in psychophysical judgment. *Perception and Psychophysics*, *1*, 439–446.

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

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.

Walker, J.T., & Scott, K.J. (1981). Auditory-visual conflicts in the perceived duration of lights, tones, and gaps. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 1327–1339.

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. (1995). Categorical scaling of stimulus duration by humans. *Journal of Experimental Psychology: Animal Behavior Processes*, *21*, 318–330.

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

Wearden, J.H., & Ferrara, A. (1995). Stimulus spacing effects in temporal bisection by humans. *Quarterly Journal of Experimental Psychology*, *48B*, 289–310.

Wearden, J.H., & Ferrara, A. (1996). Stimulus range effects in temporal bisection by humans. *Quarterly Journal of Experimental Psychology*, *49B*, 24–44.

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., Rogers, P., & Thomas, R. (1997b). Temporal bisection in humans with longer stimulus durations. *Quarterly Journal of Experimental Psychology*, *50B*, 79–94.

Wearden, J.H., & Towse, J.N. (1994). Temporal generalization in humans: Three further studies. *Behavioural Processes*, *32*, 247–264.

Wearden, J.H., Wearden, A.J., & Rabbitt, P.M.A. (1997c). 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.

*Manuscript received 10 January 1997*
*Accepted revision received 29 August 1997*

## Les sons perçus plus long que les lumières: Application d'un modèle de l'horloge interne chez les humains

Trois expériences utilisant des méthodes de généralisation temporelle et d'estimation verbale ont étudié les jugements de la duration de stimuli auditifs (des sons de 500 hz) et visuels (carré bleu de 14 cm). Avec ces deux méthodes, les stimuli auditifs furent jugés plus long et moins variable que les stimuli visuels. Les expériences sur les estimations verbales ont utilisé des stimuli variant de 77 à 1183 cm de longueur, et l'inclinaison de la fonction reliant les estimations de moyene aux vrais longueurs était differente pour les deux modalités (mais non l'intercept), ce qui concorde avec l'idée qu'une horloge interne fonctionait plus rapidement pour les stimuli auditifs que pour les stimuli visuels. Cette difference fut attribuée à une difference de variabilité dans l'opération d'un relais dans l'horloge interne, et les données expérimentales suggèrent que ces effets du relais peuvent être séparés des changements de la vitesse de l'horloge. Ces travaux montrent comment un modèle horloge concordant avec la théorie du minutage scalair ("scalar timing"), l'explication la plus reconnu du minutage animal, peut être appliquée à la perception du temps chez les humains.

# Por que ''los sonidos se juzgan mas largos que las luces'': Aplicacion de un modelo del reloj interno en humanos

Tres experimentos, en los que se usaron los metodos de generalizacion temporal y juicio verbal, estudiaron los juicios de la duracion de estimulos auditivos (tono de 500 hz) y visuales (cuadrado azul de 14 cm). Con ambos metodos, los estimulos auditivos se juzgaron mas largos, y menos variables, que los visuales. En los experimentos de juicio verbal se usaron estimulos de 77 a 1183 mseg de duracion y la pendiente de la funcion que relaciona el juicio medio con la duracion real difirio entre las modalidades (pero no la interceptada), coherente con la idea de que un marcapasos que genera representaciones de duracion corria mas rapido para los estimulos auditivos que para los visuales. La diferente variabilidad de los estimulso auditivos y visuales se atribuyo a la variabilidad diferencial en el funcionamiento de un conmutoador de un reloj marcapasos-acumulador y los datos experimentales sugerian que tales efectos del conmutador eran separables de los cambios en la velocidad del marcapasos. En conjunto, el trabajo mostro como un modelo del reloj coherente con la teoria de medida del tiempo escalar, la principal explicacion de la medida del tiempo en animales, puede aplicarse a una cuestion derivada de la literatura clasica sobre la percepcion humana del tiempo.