**Is there an “equalizing tendency” in temporal bisection in humans?**

John H. Wearden1, Kelly Bajic2 and Joanna Brooks2

1School of Psychology, Keele University, Keele, Staffordshire, ST5 5BG, U.K.,

2University of Manchester, U.K.

Email: j.h.wearden@keele.ac.uk

1ORCID 0000-0001-6589-4613

Abstract

Three experiments tested the conjecture that people carrying out a temporal bisection experiment have an “equalizing tendency”, that is, the idea that they are motivated to try to equalize the number of SHORT and LONG responses they produce. All three experiments varied instructions between two conditions. In one case, people were instructed to try to equalize, in the other it was emphasized that any proportion of SHORT and LONG responses was acceptable. Experiment 1 used short and long standards of 200 and 800 ms with comparisons spaced linearly, logarithmically, or reverse logarithmically between them, in three different groups. No effect of instructions was found in any group. Experiment 2 once again used 200 and 800 ms as the standards but had a range of comparisons closely spaced around the midpoint (500 ms), so most were ambiguous with respect to the standards. Again, no effect of instructions was found. Experiment 3 used smaller long/short ratios (300/200 ms or 800/533 ms). Once again, no effect of instructions was found. The three experiments together also replicated some previously obtained effects such as those of stimulus spacing and long/short ratio.

Key words: temporal bisection; instructions; stimulus spacing; stimulus range

The temporal bisection procedure with humans, often supposed to have been introduced into modern timing research by Allan and Gibbon (1991) and Wearden (1991), although Bovet (1968) did a much earlier study, has been very extensively used to evaluate timing performance in a variety of participant groups, including children (McCormack et al., 1999; Droit-Volet & Wearden, 2001), older people (Wearden et al., 1997; McCormack et al., 1999), and clinical groups (e.g. Wearden et al., 2008), as well as student participants. Two main versions of this task exist, one (*similarity*) initially presents standards identified as short and long (*S* and *L*, for example, tones of 200 and 800 ms duration). Then participants receive a range of comparison durations (usually including *S* and *L* but with a number of intermediate values) and are required to classify each one in terms of its similarity to *S* and *L* (or in rare cases identity to *S* or *L*, as in Allan & Gibbon, 1991), making a SHORT or LONG response. In the other method (which Wearden & Ferrara, 1995, call *partition*), no stimuli are initially identified as *S* and *L*, but the range of stimulus values is presented repeatedly and each stimulus is classified as SHORT or LONG. Using the similarity method, feedback is sometimes provided when *S* and *L* are presented as comparisons but, alternatively, identified values *S* and *L* are presented again after a number of responses to comparison stimuli.

The result of both bisection procedures (which yield similar results, see Wearden & Ferrara, 1995) is a psychophysical function generally expressed in the form of the proportion of LONG responses plotted against comparison stimulus duration. This psychophysical function is usually analyzed to derive two measures. One of these is the *bisection point* (BP), the comparison value giving rise to 50% LONG responses, and the other is a Weber-fraction-like measure, based on the slope of the psychophysical function, which is a measure of temporal sensitivity.

The location of the BP in studies with humans has given rise to much discussion. In animal studies (e.g. Church & Deluty, 1977; Meck, 1983) it is almost invariably located at the geometric mean of *S* and *L* (the square root of their product). In humans, the location is much less clear, and appears to depend on a number of factors, such as the spacing of stimuli between *S* and *L* (Wearden & Ferrara, 1995) and the *L/S* ratio (Wearden & Ferrara, 1996), with large ratios tending to produce something close to arithmetic mean bisection (that is the BP = (*S* + *L*)/2) and smaller ratios producing BPs closer to the geometric mean (which is arithmetically always smaller than the arithmetic mean). Kopec and Brody (2010) provide a thorough review, generally concluding that the BP is “sub-arithmetic”, that is, below the arithmetic mean but not necessarily always close to the geometric mean.

Some models of bisection invoke a principle that we will here call *equalizing*. This assumes that participants “make use of..an implicit assumption that half of their responses should be short and half long” (McCormack et al., 1999, p. 1152). In the model of McCormack et al. (1999), subjective durations are supposed to be logarithmically related to real duration, so by itself their model will predict geometric mean bisection, which they did not usually find, so the tendency to bisect at the geometric mean is ‘corrected’ by equalizing. Kopec and Brody’s (2010) own model of bisection used a very similar idea, proposing (p. 268) that “long and short responses should be produced with equal likelihood”. The basis for this equalizing principle was what they called the “gambler’s fallacy”: that is, in a situation with two outcomes, like heads or tails, after a run of one outcome a person increasingly expects the other one. In Kopec and Brody’s case, they proposed that a run of, say, SHORT responses would encourage a person to respond LONG, and vice versa, and their model included a bias parameter which reflected this hypothesized effect.

At first sight, one might suppose that the equalizing tendency, at least in Kopec and Brody’s form, might be measured by examining the sequence of stimuli and responses occurring in a bisection task. This seems certainly possible, but a moment’s thought suggests that this might be more difficult that it seems initially. Most bisection experiments present the comparison stimuli in a random order, so obtaining a sufficient number of instances where a number of LONG responses are followed by a SHORT one that may not seem justified on the basis of stimulus duration alone may require an unrealistic number of experimental trials from each participant. Sequences could be constructed so that stimuli that might be expected to give rise to a considerable number of consecutive SHORT or LONG responses are followed by an intermediate one, and the influence of the prior stimuli observed, but such sequences may be unrealistic and not reflect the sequences of stimuli that occur in most bisection experiments. In addition, variability in temporal representations from one trial to another further complicates observations, as does an apparent tendency for assimilation effects in bisection (Wehrman et al., 2023), that is, a tendency for a stimulus to give rise to the same response as the stimulus on the previous trial.

How then can the possibility that bisection performance in humans might be affected by an “equalizing” tendency be evaluated empirically? The present article reports three experiments which attempt to manipulate the response strategies that people use in temporal bisection by means of instructions. Essentially, all three experiments are variants of one another: in all three people perform bisection on a set of durations twice. In one case, instructions are given that encourage the participant to try to equalize the number of SHORT and LONG responses in the experimental session (equalize instructions: EQ). In the other case, people are instructed to classify each test duration individually, without regard to the overall number of SHORT and LONG responses produced (individualize instructions: IND). The obvious focus of interest is on the differences, if any, between the psychophysical functions produced in the two cases, with particular emphasis on location of the bisection point.

Although the main aim of the three experiments was to try to determine the extent to which an imposed response strategy affected bisection in humans, the three experiments overall also investigated a number of other issues in temporal bisection in humans, by providing examples of the possible effect of stimulus spacing (Experiments 1 and 2), and stimulus range (Experiment 3) on bisection performance.

## EXPERIMENT 1

In Experiment 1, three groups of participants each performed the bisection of a set of durations twice, once in each instructional condition. The different groups differed only as to the spacing of non-standard durations between constant *S* and *L* values (200 and 800 ms). In one group, the test durations were linearly spaced between *S* and *L*, in another group the spacing was logarithmic, and in the third group a “reverse logarithmic” set was used. The aim of using different stimulus spacings between a constant *S/L* pair was twofold. Firstly, we aimed to see if any instructional effect would manifest itself more clearly with one stimulus spacing than another. In the different sets, the durations were differently biased with respect to the arithmetic mean of S and *L*. In the linear conditions, there were the same number of durations above the mean as below it, and the mean of the set was also the mean of *S* and *L* (500 ms). In the logarithmic spacing group, there were more durations below the *S/L* mean than above, and the mean of the set was below the *S/L* mean. In contrast, in the “reverse” logarithmic spacing, there were more durations above the *S/L* mean than below it, and the mean of the set was also above the *S/L* mean. If the equalize instruction affected behaviour by shifting the bisection point towards the arithmetic mean of *S* and *L*, then the manipulations might well have different effects in the different spacing conditions. The second aim of varying stimulus spacing was to try to replicate the effects reported in Wearden and Ferrara (1995), who found that, relative to a linear spacing condition, logarithmic spacing shifted the psychophysical function to the left, and a spacing which biased the set towards *L* (similar to our “reverse” logarithmic condition, below), shifted it to the right. Similar results from bisection in humans have also been reported by Allan (2002).

### Method

#### Participants

48 University of Manchester Psychology undergraduates participated for course credit. They were arbitrarily allocated to 3 equal-sized groups.

#### Apparatus

The experiment took place in a small cubicle insulated from external lights and noise. An OPUS (IBM-compatible) PC with a standard black and white monitor controlled all experimental events. All stimuli to be timed were 500 hz tones produced by the computer speaker, and the keyboard was used to record responses. The experimental program was written in the MEL language (Psychology Software Tools, Inc.) thus ensuring ms accuracy for the timing of stimuli and responses.

#### Procedure

Participants received a single experimental session lasting 25-30 minutes in which they experienced both instructional conditions. Half the participants in each group performed under the equalize (EQ) instructions, then the individualize (IND) instructions; for the other half the order of conditions was reversed. The participants were given general instructions describing the task, which described the initial presentations of Short (*S*) and Long (*L*) standards, and the requirement to classify subsequent test stimuli in terms of their similarity to *S* and *L* (by pressing “S” or “L” on the keyboard). Before performing in the two conditions, the participants received the following instructions.

#### Equalize: EQ

“For this experiment, while you’re deciding whether the tones you hear are more similar to the long or short standards, I’d like you also to try to respond with about equal numbers of ‘S’ and ‘L’ answers. It’s not important that you respond exactly half ‘S’ and half ‘L’, and you shouldn’t spend too long thinking about your answer. Just keep in the back of your mind the thought that the proportions of ‘S’ and ‘L’ answers should be around equal.”

#### Individualize: IND

“For this experiment, please don’t worry at all about how many ‘S’ and ‘L’ responses you make. It doesn’t matter if you answer ‘S’ a lot more than ‘L’ or if you choose ‘L’ a lot more than ‘S’. Don’t spend too long thinking about your answer. All you need to decide is whether a tone sounds more like the short or long standard”.

After the *second* set of instructions, whatever they were, was read out, the following instructions were given.

“Although you’ve been given different instructions to follow for each experiment, please don’t think that you should necessarily be doing something different each time. We don’t know what effect the instructions will have on how you carry out the experiment – that is what we’re trying to find out!”

To start the experimental condition, the participant pressed the spacebar in response to a prompt. This produced the display “You will now receive four presentations of the standard Short and standard Long time intervals. The standard Short will come first, then the standard Long”. Presentations of *S* (200 ms) and *L* (800 ms) were separated by interpresentation intervals ranging between 1500 and 2500 ms. Participants then received the display “You will now receive a number of comparison stimuli. Press spacebar for next trial". The response was followed by a delay randomly chosen from a uniform distribution running from 2000 to 3000 ms, then one of 7 comparison stimuli. The set of comparison stimuli comprised *S* and L, and the 5 durations shown in Table 1, which differed for the different experimental groups (LIN, LOG, and REV). The order of the 7 stimuli was randomized between participants. After all 7 had been presented, the participant received 4 presentations of *S* and *L* as at the start of the experiment. The experiment continued until each comparison stimulus had been presented 10 times, so 10 judgements of the similarity of each comparison stimulus to *S* and *L* was obtained. As noted above, each participant performed the bisection task twice, once with EQ instructions, once with IND instructions, but the stimulus spacing for a particular participant was kept constant.

Table 1 and Figure 1 about here

#### **Results and Discussion**

Figure 1 shows psychophysical functions (proportion of LONG responses plotted against comparison stimulus duration) for the three stimulus spacing conditions (linear: upper panel; log: center panel; REV: bottom panel). Data were taken from blocks 4-10 of the Experiment. In each panel data from the different instructional conditions (EQ and IND) are shown separately. Inspection of the results suggests that (a) the psychophysical functions were very orderly in all cases (e.g. the comparison that was *S* produced hardly any LONG responses, and the comparison that was *L* produced nearly 100% LONG responses), (b) the effect of instructions was small with all three stimulus spacings, and (c) relative to data from the LIN spacing group, the psychophysical function from the LOG spacing group was displaced to the left, and that from the REV group displaced to the right.

We conducted an ANOVA with instruction (EQ versus IND) and comparison stimulus duration (just called duration, below) as within-group factors and stimulus spacing (LIN, LOG, REV) as the between-group factor. There was a significant between-group effect, *F*(2, 45) = 11.38, *p* < .001, a significant within-group effect of duration, *F*(6, 270) = 642.64, *p* < .001, and a significant duration x stimulus spacing group interaction, *F*(12, 270) = 6.02, *p* < .001. Effect of instruction, *F*(1, 45) = 3.20, and the instruction x stimulus spacing group interaction, *F*(2, 45) = 2.99, both approached significance, *p* = .08 and .06 respectively, but neither the instructional condition x duration interaction, *F*(6, 270) = 1.33, nor the three way interaction, *F*(12, 270) = .47, was significant.

The ANOVA indicated that the different stimulus spacings produced different proportions of LONG responses overall, and that the stimulus spacing effect was different at the different comparison durations (i.e. the psychophysical curves with different stimulus spacings were displaced relative to one another). A final, and obvious, effect was that different stimulus durations gave rise to different proportions of LONG responses

Table 2 about here

Table 2 shows bisection points and Weber ratios calculated from the different experimental conditions. To calculate the *bisection point*, a two-parameter logistic function was fitted to the psychophysical functions from the different conditions (as in Morrisey et al., 1993), using the non-linear regression subprogram of SPSS 10. The equation fitted was

*y* = 1/(1 + (*t/b*)*s* )

where *y* is the proportion of LONG responses made to a stimulus duration *t*, and *b* and *s* are fitting constants. *b* is the bisection point, the stimulus duration giving rise to 50% LONG responses, and *s* is a measure of the slope of the psychophysical function. Fits of the logistic equation had *r2* values of at least .99.

Another measure of interest is the *Weber ratio*, an index of temporal sensitivity that essentially reflects the steepness of the psychophysical function, where smaller values indicate higher temporal sensitivity. The logistic equation was used to calculate the duration values that would give rise to 25% and 75% LONG responses. Half the difference between these values, divided by the bisection point for the condition, is the Weber ratio.

Inspection of the values in Table 2 suggests that stimulus spacing had a marked effect on bisection point (consistent with the significant displacements of the psychophysical functions mentioned above). The bisection point from the LOG group was below that from the LIN group, whereas the REV manipulation increased bisection point relative to the LIN group. A fuller discussion of the bisection point location will be presented later, but we note here that in the LIN and REV groups the bisection point from both instructional conditions was closer to the arithmetic mean of *S* and *L* (500 ms) than the geometric mean (400 ms), whereas for the LOG group the reverse was the case. Weber ratios remained within a very narrow range (0.12 to 0.15) whatever the stimulus spacing and instructional condition.

The bisection points in Table 2 were used to test the data from Experiment 1 for the property of *superposition*, the requirement that data from different timed durations superpose when plotted on the same relative scale, a strong requirement of scalar timing theory (Gibbon, Church, & Meck, 1984). To examine superposition, the proportion of LONG responses at each comparison duration is plotted against the comparison duration expressed as a fraction of the bisection point for the condition in force. Figure 2 shows the results when this was done. Inspection of Figure 2 shows that superposition was almost perfect, so the scalar property of time representations held in our data regardless of effects of stimulus spacing and instructions.

Figure 2 about here

Overall, therefore, our data replicated some previous findings by Wearden and Ferrara (1995) that spacing of comparison durations between a consistent *S/L* pair significantly affected the psychophysical functions, with a LOG spacing displacing the curve to the left relative to LIN, whereas the REV manipulation (and the similar manipulation in Wearden and Ferrara, 1995, Experiment 2) displaced it to the right. In contrast to the effect of stimulus spacing, the effect of instructions was very small, and not statistically significant. This suggests that using an “equalize” or “individualize” response strategy in temporal bisection had little or no effect on overall outcome.

Why was the instructional effect in our study so weak? One possibility is that instructions, and the attendant response strategies they engender, have effects on bisection only in certain conditions. For example, with a 200/800 ms *S/L* pair, many comparison durations are “unambiguous” to our participants, in the sense that with the discriminative acuity of young adults, some stimuli are “obviously” closer to *S* than to *L*, and vice versa. An instructional manipulation cannot affect these judgements without violating the demand characteristics of the experiment, as any instruction which attempted to encourage participants to respond LONG to short comparison durations, or SHORT to long comparisons would require the participants to violate their perceptual experience, and thus conflict with the general task instructions for bisection. To give a fanciful example, suppose that *S/L* values were 200/800 ms, but that all comparison durations were between 200 and 300 ms, or between 700 and 800 ms. A participant could not obey the “equalizing” instructions without violating their perceptual experience, and the general instruction to classify durations in terms of their similarity to *S* and *L*.

Following this argument, one might expect more marked effects of instructions on response strategies if comparison durations were more ambiguous relative to *S* and *L* (for example, being somewhere in the middle of the *S* to *L* range, rather than toward either end). Experiment 2 tested this idea by employing an *S/L* pair of 200/800 ms, but arranging that the majority of the comparison durations were located in the middle of the *S* to *L* range.

## EXPERIMENT 2

#### **Method**

#### Participants

15 Manchester University Psychology undergraduates participated for course credit.

#### Apparatus

As Experiment 1.

#### Procedure

The procedure was identical to that of Experiment 1 in all particulars except for two differences. The comparison durations ranged from 200 to 800 ms but, as shown in Table 1, intermediate durations were not equally spaced, so that most of 9 them were close to 500 ms. Apart from *S* and *L* themselves, comparison durations ranged from 350 to 650 ms, in 50-ms steps. 7 blocks of these 11 comparison durations were given, rather than 10 as in Experiment 1.

#### **Results and Discussion**

Figure 3 shows the psychophysical functions from the EQ and IND conditions of Experiment 2, with data coming from blocks 3-7 of the Experiment. Inspection of the Figure suggests that the psychophysical functions were orderly in both instructional conditions, with minimal effect of instructions. These suggestions were confirmed by statistical analysis. There was no significant effect of instructional condition, *F*(1, 14) = .16, *p* = .70, nor of instruction x duration interaction, *F*(10, 140) = .33, *p* = .97, but there was the usual highly significant effect of duration, *F*(10, 140) = 69.97, *p* < .001. Table 2 shows bisection points and Weber ratios derived in the same manner as for Experiment 1. Fits of the logistic equation produced *r2* values of at least .99. The bisection points from the two instructional conditions were both closer to the arithmetic mean of *S* and *L* than the geometric mean, and Weber ratios were close to those obtained in Experiment 1.

Figure 3 about here

Experiment 2 suggests that changing response strategies in bisection had no effect on performance, even though many of the comparison durations used in the experiment were mid-way between *S* and *L*, and were thus likely to be ambiguous with respect to the SHORT or LONG standards.

Experiment 2, by using comparison stimulus durations in the mid-range of *S* and *L* employed one method of increasing the ambiguity of comparison durations. Another is to make the *L/S* ratio much smaller. The effect of *L/S* ratio on bisection found in previous studies (Wearden & Ferrara, 1996; Ferrara et al., 1997) will be discussed in more detail later, but it may be that the responding of our participants was insensitive to instructional manipulations because a large *L/S* ratio (4) was used. In Experiment 3, two groups of participants performed bisection twice, with EQ and IND instructions, as above, but this time the different groups bisected sets with a *L/S* ratio of 1.5:1. Both stimulus sets came from the 200-800 ms range, to facilitate comparison with the data obtained in Experiments 1 and 2, but differed in that the two groups (*Short* and *Long*) came from opposite ends of the range. For the short group *S* was 200 ms, and *L* 300 ms; for the long group *S* was 533 ms and *L* 800 ms. In both cases comparison durations were spaced linearly between *S* and *L*.

## EXPERIMENT 3

### Method

*Participants*

30 undergraduate students were allocated to two equal-sized groups.

*Apparatus*

As Experiments 1 and 2.

*Procedure*

The procedure was identical to that used in Experiments 1 and 2 except for the stimulus durations used . For the *Short* group *S* = 200 ms and *L* = 300 ms; for the *Long* group, *S* = 533 ms and *L* = 800 ms. For both groups the intermediate comparison stimuli were linearly spaced between *S* and *L*, as shown in Table 1.

### Results and Discussion

The upper panel of Figure 4 shows mean proportion of LONG responses plotted against comparison duration from the *Short* group, and the center panel the same measure from the *Long* group, with data coming from blocks 4-10 of the experiment. One participant in the *Long* group produced an abnormally smaIl proportion of LONG responses to *L* (.29, with the next smallest in the group being over 0.8), so all data from this participant were excluded from the analysis. Inspection of the data suggests that the psychophysical functions were orderly in both groups, ranging from virtually zero LONG responses when *S* was presented, to over 90% when *L* was presented. In neither group did there appear to be a clear effect of instructional condition (EQ versus IND).

Figure 4 about here

Consider first data from the *Short* group. ANOVA found no effect of instructions (EQ versus IND), *F*(1, 14) = 2.46, *p* = .14, nor instructions x comparison duration interaction, *F*(6, 84) = .47, *p* = .83, but there was a significant effect of comparison duration, *F*(6, 84) = 147.49, *p* < .001. The bisection point and Weber ratio were calculated as for the previous experiments, and results are shown in Table 2. .The logistic function produced r2 values of at least .98 in all conditions. For both the EQ and IND conditions, the bisection point was within 1 ms of the arithmetic mean of *S* and *L* (250 ms), and the Weber ratio for both conditions was identical.

For the *Long* group (centre panel of Figure 4), there was no significant effect of instructional condition (EQ versus IND), *F*(1,13) = .452, but there was a significant effect of comparison duration, *F*(6, 78) = 157.33, *p* < .001. In addition, there was a significant instructional condition x duration interaction, *F*(6, 78) = 4.90, *p* < .01. This resulted from a between-condition difference when the comparison duration was 667 ms, *t* (13) = 3.04, *p* < .01, although it is unclear why this difference occurred. Bisection points for both EQ and IND conditions at most about 1% away from the arithmetic mean of *S* and *L*, and Weber ratios from the two conditions were almost identical.

The bisection points were used to test data from the conditions of the *Short* and *Long* groups for superposition, by plotting the proportion of LONG responses against comparison duration/bisection point, with the bisection point values coming from Table 2. The bottom panel of Figure 4 shows the results when this is done.

Inspection of the panel suggests that superposition was good, with no set of data points appearing to deviate from any other. Such good superposition would also be expected from the closeness of the Weber ratios from all the conditions of Experiment 3, with all values being either 0.08 or 0.07.

Overall, therefore, the results of Experiment 3 confirmed the general pattern of data noted in the other studies. Even though *L/S* ratio was smaller than in the previous Experiments, the discrimination between *S* and *L* presumably harder, and intermediate comparison durations close to both *S* and *L*, this did not produce any marked instructional difference: People bisected the duration set in an orderly way, with a bisection point virtually identical to the arithmetic mean of *S* and *L*, as in the linear spacing conditions from Experiments 1 and 2. Data from Experiment 3 also conformed well to the principle of superposition (Figure 4).

## GENERAL DISCUSSION

 The three experiments in the present article have unanimously supported the view that temporal bisection performance in student-age participants was little affected by the instructional manipulations employed. Whether participants attempted to equalize the number of LONG and SHORT responses they made, or treated each comparison duration individually, results were generally almost identical. The lack of significant effect of instructions found in Experiment 1 was confirmed in Experiment 2, where most comparison durations were close to the mid-range of *S* and *L*, and in Experiment 3, where the *L/S* ratio was small. It seems, therefore, that differences in bisection point location between humans and animals, or differences in bisection point location between different studies with humans, are unlikely to be caused by the adoption of an “equalizing” response strategy, contrary to the suggestion of McCormack et al. (1999) and Kopec and Brody (2010).

It should be emphasized that the data in Experiments 1 to 3 above appear entirely “normal” in the sense that they replicated many results obtained in previous studies of temporal bisection in humans. One notable result was the location of the bisection point. When comparison durations were linearly spaced between *S* and *L* this was close to the arithmetic mean of *S* and *L* in all experiments, as found in other studies (Kopec & Brody, 2010). Another effect was that changing the spacing of comparison durations between a constant *S/L* pair affected bisection point location: logarithmic spacing lowered the bisection point relative to linear spacing, whereas “reverse log” spacing increased it. This effect of stimulus spacing replicates that originally found by Wearden and Ferrara (1995), and replicated by Allan (2002). For example, Allan’s (2002) Table 2 (p. 54, but the reader should note that there is a mislabelling in the Table and “1000lin” should be “600log”) provides 15 pairs of bisection points from linear and logarithmic spacing between fixed *S/L* pairs, and in 14 cases the bisection point from the logarithmic case was lower. Experiment 1 also replicated the result from Wearden and Ferrara (1995) that duration sets with the opposite skew to a logarithmic set (i.e. more values above the midpoint of *S* and *L* than below it) shift bisection points to the right of those obtained with linear spacing.

Another effect obtained in our data was a duplication of the finding of Ferrara et al. (1997) that more “difficult” bisection tasks, with difficulty assayed by the size of the *L/S* ratio, produced more sensitive timing in terms of smaller Weber ratios. Experiments 1 and 2, above, used a 4:1 *L/S* ratio, whereas Experiment 3 employed a 1.5:1 ratio, and the Weber ratio decreased accordingly.

Data like those collected here can be modelled fairly accurately by a simple model which does not assume any equalizing principle. We used a slight modification of the Wearden and Ferrara (1995) model which, in its original form, used the (unadjusted) arithmetic mean of the set of comparison durations as the criterion and the source of variance in the model, with the comparison durations being timed without error. We modified the model simply assuming that 95% of the arithmetic mean (95%AM) was used as the criterion. Justification for this is provided below. So, if *M* is 95%AM, *t* is the duration to be judged, and *b* a threshold, the model responds LONG when (*t* - *M*)/*t* > *b*, responds SHORT when (*t* - *M*)/*t* < -*b*, and when -*b* < (*t* - *M*)/*t* < *b*, the model makes LONG and SHORT responses at random and with equal probability (see Wearden and Ferrara, 1995, p. 305, for further discussion).

The model has two parameters, *b*, the threshold, and *c*, the coefficient of variation [Weber fraction] of the memory of the criterion, *M*. Figure 5 shows simulations of bisection performance using the durations appropriate to the LIN, LOG, and REV groups of Experiment 1. Figure 5 shows averaged results from 10000 simulated trials with each comparison duration, as well as data points from Experiment 1, with data from the EQ and IND conditions of the LIN, LOG, and REV groups being averaged together to produce the data points shown in the Figure (as unconnected symbols). In the simulations whose results are shown, 95%AM was used as *M*, the Weber ratios from Table 2 for the appropriate conditions (.13, .14, and .13 for the LIN, LOG, and REV conditions, respectively) were used as *c*, and the threshold was varied. *M* was represented as a Gaussian distribution with accurate mean and coefficient of variation, *c*, and a random value chosen from this distribution served as the criterion on each trial.

Figure 5 about here

Consider first the lines without symbols in Figure 1, the results of modelling LIN, LOG, and REV conditions. It is clear that (a) the psychophysical functions predicted are very orderly, with the proportion of LONG responses varying from near zero when *S* is presented, to near 100% when *L* is presented, and (b) that the effect of stimulus spacing (LIN, LOG, REV) emerges naturally from the model. Comparison of the theoretical curves and the data suggests that the two are close together (in fact, mean absolute deviation was .02, .02, and .01, for the LIN, LOG, and REV groups, respectively). This closeness is all the more remarkable as no attempt has been made to adjust all the parameters of the model to fit data from the groups individually: that is, we always used 95%AM (and not some similar value, which could differ for the different groups), and the coefficient of variation used was the averaged Weber ratio from Experiment 1 for the appropriate condition. The only parameter varied during the modelling was *b*, as this took values of either .15 or .14 in all cases.

The obvious question that arises is why 95% of the arithmetic mean of the set of comparison stimuli and not 100%, or some other value? A plausible possibility perhaps comes from considering an asymmetry between the shorter and longer durations in terms of *when* the decision to respond SHORT or LONG can be made. The shorter durations in the set of stimuli used for the bisection experiment cannot be classified as SHORT until they have terminated as, of course, it is possible that they would have continued for longer than they did. On the other hand, the longer durations, particularly the very longest ones, may be classified as LONG when they have perceptibly exceeded the criterion in force, which can happen *before* they have actually terminated. One possibility is that the participant registers these longer durations as slightly shorter than they really are, perhaps ceasing to accumulate elapsing time once the decision to respond LONG has been made. A tendency to do this would subjectively shorten the longer durations slightly (whereas the shorter ones would always be registered as their full value), and if these slightly shortened values are aggregated with others to produce some mean time value used as a criterion, the mean used will be slightly shorter than the arithmetic mean. Indirect support for this idea comes from Klapproth and Wearden (2011) who used a temporal generalization experiment, where people have to judge whether a comparison stimulus has the same duration as a previously-presented standard. In some conditions, people were asked to make the response as quickly as possible, even before the stimulus had terminated, and response time data suggested that the decision about the response to the longer stimuli presented could occur before the stimulus had finished, for the longer stimulus durations.

It is possible that some equalizing tendency in bisection might be observed in some as yet untested conditions in spite of our persistent failure to find it evidence for it in our experiments. Obviously, there are logical problems proving a negative, but our experiments (particularly experiment 2 with many “ambiguous” comparison stimuli) seem to have given the equalizing tendency considerable opportunity to manifest itself if it was present and there was no instructional effect. It is worth noting that the equalizing tendency is a feature needed in theoretical models in both McCormack et al. (1999) and Kopec and Brody (2010) to enable them to accurately simulate data, rather than something actually observed in previous experiments, so our failure to obtain an effect of equalizing instructions does not conflict with any existing body of experimental data.

Table 1

Stimulus durations employed in Experiments 1 to 3. **S** = short standard, **L** = Long standard. All values are in ms.

-----------------------------------------------------------------------------------

 **S L Intermediate durations (ms)**

**Experiment 1**

LIN 200 800 300,400,500,600,700

LOG 200 800 252,317,400,504,635

REV 200 800 365,496,600,683,748

# **Experiment 2**

 200 800 350,400,450,475,500,525,550,600,650

# **Experiment 3**

*Short*  200 300 216,233,230,267,284

*Long*  533 800 577,622,667,711,756

Table 2

Bisection points (**BP**, in ms) and Weber ratios (**WR**) from the different experimental conditions of Experiments 1, 2, and 3.

------------------------------------------------------------------------------------------

###### BP WR

##### Experiment 1

LIN(EQ) 482 .12

LIN(IND) 451 .13

LOG(EQ) 421 .15

LOG(IND) 424 .13

REV(EQ) 536 .12

REV(IND) 515 .13

##### Experiment 2

EQ 486 .12

IND 480 .14

**Experiment 3**

*Short* (EQ) 250 .07

*Short* (IND) 251 .07

*Long* (EQ) 659 .08

*Long*(IND) 692 .07

**REFERENCES**

Allan, L.G. (2002). The location and interpretation of the bisection point. *Quarterly Journal of Experimental Psychology, 55B*, 43-60.

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

Bovet, P. (1968). Echelles subjectives de durée obtenues par une méthode de bisection. *L’Année Psychologique*, *68*, 23-36.

Church, R.M., & Deluty, M.Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology: Animal Behavior Processes*, *3*, 216-228.

Droit-Volet, S., & Wearden, J.H. (2001). Temporal bisection in children. *Journal of Experimental Child Psychology*, *80*, 142-159.

Ferrara, A., Lejeune, H., & Wearden, J. H. (1997). Changing sensitivity to duration in human scalar timing: An experiment, a review, and some possible explanations. *Quarterly Journal of Experimental Psychology*, *50B*, 217-237.

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*, 423: Timing and time perception, (pp. 52-77). New York: New York Academy of Sciences.

Klapproth, F., & Wearden, J.H. (2011). Why do temporal generalization gradients change when people make decisions as quickly as possible? *Quarterly Journal of Experimental Psychology*, *64*, 1646-1664.

Kopec, C.D., & Brody, C.D. (2010). Human performance on the temporal bisection task. *Brain and Cognition*, *74*, 262-272.

McCormack, T., Brown, G.D.A., Maylor, E.A., Darby, A. & Green, D. (1999). Developmental changes in time estimation: Comparing childhood and old age. *Developmental Psychology*, *35*, 1143-1155.

Meck, W.H. (1983). Selective adjustment of the speed of internal clock and memory processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 171-201.

Morrissey, G., Wogar, M.A., Bradshaw, C.M., & Szabadi, E. (1993). Effect of lesions of the ascending 5-hydroxytryptaminergic pathways on timing behaviour investigated with an interval bisection task. *Psychopharmacology*, *112*, 80-85.

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

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

Wearden, J.H., Smith-Spark, J.H., Cousins, R., Edelstyn, N.M.J., Cody, F.W.J., & O’Boyle, D.J. (2008). Stimulus timing in people with Parkinson’s disease. *Brain and Cognition*, *67*, 264-279.

Wehrman, J., Sanders, R., & Wearden, J. (2023) What came before: Assimilation effects in the categorization of time. *Cognition*, *234*, 105378.

**Figure legends**

Figure 1. Psychophysical functions from the LIN (upper panel), LOG (centre panel), and REV (lowest panel) groups of Experiment 1. Data are plotted in terms of the mean proportion of LONG responses (i.e. judgements that a presented duration was more similar to the Long standard than the Short one) against stimulus duration. In each panel, data from the EQ and IND instructional conditions are shown.

Figure 2. Mean proportion of LONG responses from the different groups and instructional conditions of Experiment 1, plotted against stimulus duration divided by the bisection point for the conditions plotted. Bisection points are shown in Table 2.

Figure 3. Psychophysical functions from the EQ and IND conditions of Experiment 2. All relevant details as Figure 1.

Figure 4. Data points from the three conditions of Experiment 1 (open symbols), and predictions of the model described in the text (lines).