A bias-free test of human temporal bisection: Evidence against bisection at the arithmetic mean

David J. Sanderson[*]

Department of Psychology, Durham University, South Road, Durham DH1 3LE, UK

ARTICLE INFO

ABSTRACT

The temporal bisection procedure has been used to assess theories of time perception. A problem with the procedure for measuring the perceived midpoint of two durations is that the spacing of probe durations affects the length of the bisection point. Linear spacing results in longer bisection points closer to the arithmetic mean of the durations than logarithmic spacing. In three experiments, the influence of probe duration distribution was avoided by presenting a single probe duration of either the arithmetic or geometric mean of the trained durations. It was found that the number of participants that categorised the arithmetic mean as long was significantly larger than those that categorised it as short. The number of participants that categorised the geometric mean as either short or long did not significantly differ. This was true for trained durations of 0.4 s vs. 1.6 s (Experiments 1–3), 0.2 s vs. 3.2 s (Experiment 2) and 0.4 s vs. 6.4 s (Experiment 3). In Experiment 4, the probe trial distribution effect was replicated with logarithmic and linearly distributed probe durations, demonstrating that bisection occurs close to the arithmetic mean with linearly spaced probe durations. The results provide evidence against bisection at the arithmetic mean when probe spacing bias is avoided and, instead, the results are consistent with logarithmic encoding of time, or a comparison rule based on relative rather than absolute differences.

1. Introduction

In a typical temporal bisection procedure, subjects are trained to discriminate between two durations of a presentation of a stimulus by making different responses for when the presentation is a short duration and a long duration (see Kopec & Brody, 2010; Penney & Cheng, 2018 for reviews). In experiments with humans, participants receive feedback on their responses. With animals, they receive reinforcement for making the correct response on a given trial. After the discrimination is acquired, subjects are presented with probe durations of the stimulus that are intermediate to the trained short and long durations and no feedback or reinforcement is presented. Categorisation of these probe durations as either 'short' or 'long', as indicated by the appropriate response allows a psychometric function to be plotted that shows the transition from making mostly 'short' responses to mostly 'long' responses as a function of time (see Fig. 1, panel a). The timepoint at which a subject makes equal 'short' and 'long' responses is the bisection point and may be considered as the perceived midpoint of the two trained durations and, therefore, the duration that is equally similar to the trained short and long durations.

Identification of the bisection point may allow the relation between subjective time and physical time to be identified. If perception of time is linear then the perceived midpoint between two durations may be their arithmetic mean (see Fig. 1, panel b). This outcome requires that the comparison of intervals reflects their difference such that the midpoint is the duration that is equally far from the short and long durations. It is possible, however, that alternative methods of comparison may lead to bisection at durations different from the arithmetic mean. For example, scalar expectancy theory proposed that time is perceived linearly, but comparison of durations is achieved by their relative rather than absolute differences (Gibbon, 1977, 1981). This results in the bisection point being the geometric mean of the durations, i.e., the square root of the product of the short and long durations. Thus, the geometric mean is the duration that is a ratio of the short duration that is the square root of the long-to-short ratio. For example, for short and long durations that have a 1:4 ratio, the geometric mean is the duration that is twice the short duration and half the long duration. If perception of time is logarithmic (see Fig. 1, panel c) then the perceived midpoint will be the geometric mean when the difference of the log durations is used. Regardless of the method for comparing the durations, logarithmic perception of time

* Corresponding author.
  E-mail address: david.sanderson@durham.ac.uk.

results in bisection below the arithmetic mean.

Experiments in animals have typically shown bisection close to the geometric mean (e.g., Church & Deluty, 1977; Platt & Davis, 1983). This has been taken as evidence for logarithmic perception of time (Church & Deluty, 1977) or, instead, evidence of comparison of relative rather than absolute differences (Gibbon, 1981). This fits in with other findings that suggest that behaviour in animals reflects sensitivity to relative rather than absolute durations, such as the time-scale invariance of conditioned response timing (Gibbon, 1977) and is in keeping with Weber's law that the discriminability of two stimuli increases as a function of their ratio rather than absolute values. The implication of this is that the discrimination between two durations with a fixed difference between the short and long duration becomes increasingly difficult as the durations increase (e.g., it is harder to discriminate between 11 s and 12 s than 1 s and 2 s despite the difference between the pairs of the durations being the same). In this respect, time perception in animals is similar to perception in other modalities in which discrimination reflects the relative rather than absolute difference between stimuli (Hecht, 1934; Moore & Raab, 1975).

In contrast to animals, in humans, it has typically been found that bisection occurs close to the arithmetic mean (e.g., Wearden, 1991; Wearden & Ferrara, 1995; Wearden, Rogers, & Thomas, 1997). This suggests that temporal perception is linear and discrimination occurs by means of comparison of the differences between durations. This result is at odds with other findings that have suggested that human time perception shows scalar properties conforming to Weber's law (Haigh, Apthorp, & Bizo, 2021; Rakitin et al., 1998).

A problem with interpretation of the bisection point as the perceived midpoint of the durations is that the bisection point is affected by the spacing of the probe durations used in the test phase (Brown, McCormack, Smith, & Stewart, 2005; Penney, Brown, & Wong, 2014; Wearden et al., 1997; Wearden & Ferrara, 1995; Zhu, Baykan, Muller, & Shi, 2021). When the probe durations are logarithmically spaced, such that there are more probe durations below the arithmetic mean than above, the bisection point is a shorter duration than when the probe durations are linearly spaced (with an equal number of probe durations above and below the arithmetic mean). Although the majority of the evidence for the stimulus spacing effect comes from studies in humans, animals are also affected by the manipulation (Raslear, 1983, 1985). These findings suggest that throughout the course of the test phase subjects are biased by the distribution of probe durations such that they come to distribute their responses equally below and above the median probe duration. Indeed, bisection is influenced by the frequency of individual probe durations, not just the spacing of durations, such that bisection points

will be shorter with higher frequency short than long probe durations suggesting that people weight their responses based on the overall statistics of the probe durations (Zhu et al., 2021). The issue of the distribution of probe durations during the test phase on interpretation of the bisection point has led to the conclusion that the bisection point cannot be taken as a pure measure of the perceived midpoint of two durations (Allan, 2002).

A way in which the effect of probe duration spacing on the bisection point can be avoided is simply to not have a distribution. This can be achieved by having a single probe duration trial. For the first probe trial of a test phase, there is no expectation of the spacing or frequency of probe durations, therefore, the bias to make a 'short' or 'long' response, independent of the perceived temporal midpoint, is zero. Performance on the first probe trial is a bias-free measure of temporal categorisation in the bisection procedure. The obvious drawback to only considering data from the first probe trial is that it would reveal nothing about temporal bisection in individuals. In a sample of participants, however, it would, at the very least, allow falsification of hypothesised bisection points.

The purpose of the current set of experiments was to test temporal bisection using a single probe trial and, thus, avoid the bias of probe duration distribution. In Experiments 1–3, participants were trained with a short and long duration and then presented with either the geometric or arithmetic mean of the trained durations. A difference in the proportion of participants responding 'short' or 'long' across the two means would provide evidence against bisection at either the geometric or arithmetic mean. In Experiment 1, participants were trained with 0.4 s as the short duration and 1.6 s as the long duration. The trained short and long durations were 0.2 s and 3.2 s in Experiment 2 and 0.4 s and 6.4 s in Experiment 3. Experiments 2 and 3 also included a group of participants trained with 0.4 s and 1.6 s as a positive control. Experiment 4 verified the effect of probe trial distribution on the bisection point using the same training procedure as Experiments 1–3. All experiments were conducted online using Pavlovia.

2. Experiment 1

In Experiment 1, participants were trained with a short, 0.4 s and long, 1.6 s presentation of a stimulus before receiving a single probe trial of the geometric or arithmetic mean of the durations. The short and long training durations fall within the range that typically lead to bisection at the arithmetic mean (Wearden & Ferrara, 1996) and have been used in experiments that have found stimulus spacing effects (Zhu et al., 2021). If participants bisect at the arithmetic mean when the effect of probe
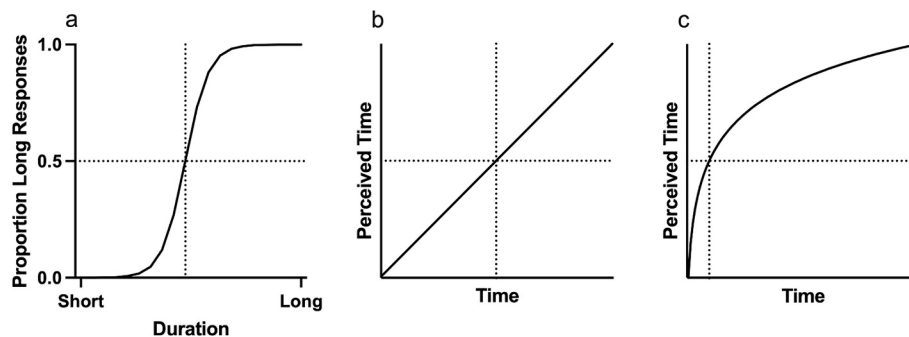

Fig. 1. Illustration of temporal categorisation of durations and the implications of linear and logarithmic perception of time. Panel a shows a hypothetical psychophysical function of proportion of 'long' duration categorisations in a bisection task as a function of probe duration. The point at which the sigmoid function crosses 0.5 on the Y axis is the bisection point, indicating the duration for which a participant makes an equal number of 'short' and 'long' categorisation responses. Panel b shows perceived time plotted as a linear function of physical time (solid black line). The dotted lines on the X and Y axes indicate the midpoint of perceived time and its relation to physical time. Panel c shows perceived time as a logarithmic function of physical time (solid black line). In contrast to linear perception (panel b), logarithmic perception of time results in the midpoint of perceived time (dotted line on the Y axis) being a shorter physical duration (dotted line on the X axis). When discrimination is based on the difference between subjective durations, the midpoint of two durations on a linear scale is the arithmetic mean and the midpoint of two durations on a logarithmic scale is the geometric mean.

duration spacing is avoided then there should be a preference for responding 'short' over 'long' when presented with the geometric mean and no preference when presented with the arithmetic mean. If participants, however, bisect at the geometric mean, there should be a preference for responding 'long' over 'short' when presented with the arithmetic mean and no preference when presented with the geometric mean.

2.1. Methods

2.1.1. Participants

One hundred and twenty-five participants were recruited from the Durham University Department of Psychology participant pool and received participant pool credit. Eleven were male, 103 were female and 11 did not report their sex. Of the 115 participants that reported their age, the mean age was 20 years (range: 17–25). The sample size for Experiment 1 and for the subsequent experiments reflects the number of participants that volunteered over an approximate four-to-six-week period. All procedures were approved by the Durham University Department of Psychology Ethics committee (PSYCH-2021-06-23T16_49_56-pggt56). Participants gave their informed consent prior to the start of the experiment.

2.1.2. Apparatus and stimuli

The discrimination task was created using Psychopy (Peirce et al., 2019) and made available online using Pavlovia. The visual stimulus used for the temporal discrimination was the presentation of a white square in the centre of the screen on a grey background.

2.1.3. Procedure

Participants received discrimination training trials in which the white square was presented for either 0.4 s or 1.6 s. There were ten trials of each duration, presented in a random order with the constraint that there was one trial of each duration every two trials. Participants were instructed that they would be presented with a temporal categorisation task in which they would learn by trial and error by receiving feedback on their responses. At the termination of the presentation of the white square participants were instructed to choose between pressing the 'Z' or 'M' key. As soon as they made a response, feedback was provided (either "Correct" or "Oops! That was wrong") for 1 s. The next trial started immediately afterwards. Pressing the 'Z' key was the correct response for the 0.4 s duration and pressing the 'M' key was correct for the 1.6 s duration.

On the 21st trial, the white square was presented for a probe duration (either the geometric mean of 0.8 s or the arithmetic mean of 1.0 s). The probe duration for each participant was randomly selected. No feedback was presented. The experiment ended when the participant made a response.

2.1.4. Data analyses

The exposition of the statistical analyses is included in the Results section. Analyses were conducted using SPSS version 29 and JASP version 0.17.2.1 (JASP-Team, 2024).

2.2. Results and discussion

An inclusion criterion of >50% correct during training was applied such that participants that failed to show performance above chance level were excluded. This resulted in the exclusion of the data from 17 participants. The mean performance of the remaining 108 participants was 87% correct, SEM = 1.2.

The distribution of 'short' and 'long' responses to the geometric and arithmetic mean probe durations is shown in Fig. 2. Three quarters of participants presented with the arithmetic mean responded 'long', whereas similar numbers of participants presented with the geometric mean responded 'long' and 'short'. A Chi square analysis of the
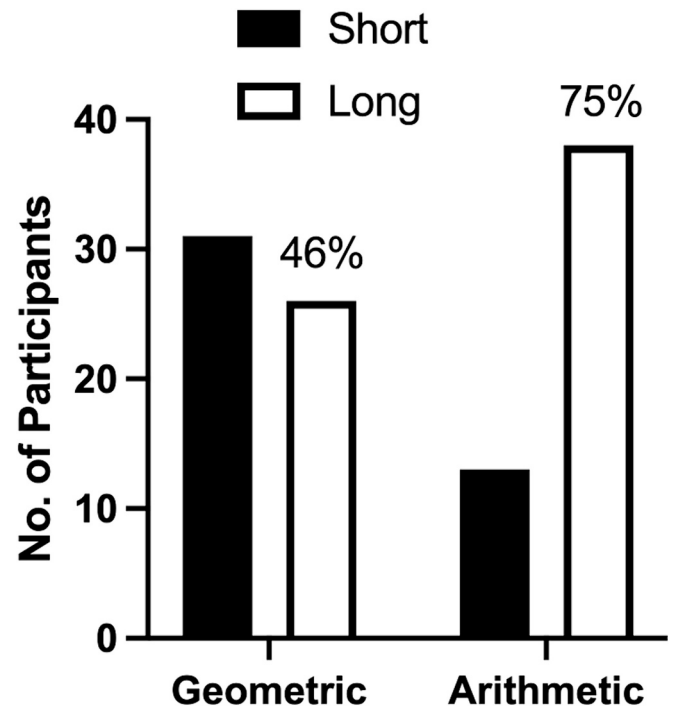

Fig. 2. The number of participants that made a 'short' or 'long' response in the probe test of Experiment 1. The percentages reported above white bars indicates the proportion of participants that responded 'long'.

distribution of responses was significant, $\chi^2 (108) = 9.31$, $p = 0.002$, $\varphi = 0.29$.

Separate binomial tests were conducted for the arithmetic and geometric groups. It was found that a significant proportion of participants responded 'long' when presented with the arithmetic mean, $p < 0.001$. The proportion of participants that responded 'long' when presented with the geometric mean was not significant, $p = 0.56$.

The results provide evidence against bisection at the arithmetic mean. The majority of participants categorised the arithmetic mean as long, but there was no significant preference for categorising the geometric mean as short or long. This suggests that the bisection point is closer to the geometric mean than the arithmetic mean. The result is in contrast to other experiments that have used similar time ranges that have found evidence for bisection at the arithmetic mean (Wearden & Ferrara, 1996). The difference between results may imply that evidence for the arithmetic mean reflects a bias as a result of the effect of the distribution of probe trials on temporal categorisation.

3. Experiment 2

A possible explanation for why Experiment 1 found evidence against bisection at the arithmetic mean may be that that the stimulus range was too short and the ratio of the short-to-long durations was too small. Although Wearden and Ferrara (1996) found evidence that durations with a 1:4 ratio (like the durations used in Experiment 1) or greater led to bisection at the arithmetic mean, ratios of 1:2 led to bisection at the geometric mean. Furthermore, evidence for the geometric mean with small short-to-long duration ratios is found only with short stimulus ranges, because longer durations (i.e., 2 s versus 4 s or 4 s versus 8 s) have resulted in bisection at the arithmetic mean (Wearden et al., 1997).

Experiment 2 replicated the procedure of Experiment 1 with the 0.4 s and 1.6 s training durations with the addition of another condition in which the short and long durations were 0.2 s and 3.2 s. Across the two conditions, the geometric mean of the intervals was matched (i.e., 0.8 s for each group) but the ratio of long-to-short durations differed fourfold. If the bisection point shifts towards the arithmetic mean as the

range between the temporal values increases, then it would be predicted that there would be a preference for responding 'short' to the geometric mean for the 0.2 s vs. 3.2 s group. If, however, bisection occurs at the geometric mean regardless of stimulus range then both groups should show a preference for responding 'long' to the arithmetic mean and the preference should be greater in the 0.2 s vs. 3.2 s group than 0.4 s vs. 1.6 s group.

3.1. Method

3.1.1. Participants

132 participants were recruited in the same manner as Experiment 1. Thirty were male, 71 were female and one person did not report their sex. Of the 131 participants that reported their age, the mean age was 20 years (range: 18–35).

3.1.2. Apparatus and stimuli

The apparatus and stimuli were the same as Experiment 1.

3.1.3. Procedure

Participants were randomly allocated to either the 0.4 s vs. 1.6 s condition or the 0.2 s vs. 3.2 s condition. In the probe trial, participants in the 0.4 s vs. 1.6 s condition were randomly allocated either the geometric mean of 0.8 s or the arithmetic mean of 1.0 s. In the 0.2 s vs. 3.2 s condition participants were randomly allocated the geometric mean of 0.8 s or the arithmetic mean of 1.7 s. All other details were the same as Experiment 1 except that for approximately half of the participants, pressing 'Z' was the correct response for the short duration and pressing 'M' was the correct response for the long duration. The opposite was true for the remaining participants.

3.2. Results and discussion

Eleven participants were excluded for failing to perform above 50% during training. The mean performance of the remaining 121 participants was 89% correct, SEM = 0.9.

The distribution of 'short' and 'long' responses to the geometric and arithmetic mean probe durations is shown in Fig. 3. For both the 0.4 s vs. 1.6 s group and the 0.2 s vs. 3.2 s group, participants predominantly responded 'long' when presented with the arithmetic mean. This was not the case for participants presented with the geometric mean of the durations.

Four separate Chi square analyses were conducted to compare the distribution of 'short' and 'long' responses amongst the groups. The first two comparisons were similar to those conducted for the results of Experiment 1, in which responses to geometric and arithmetic means
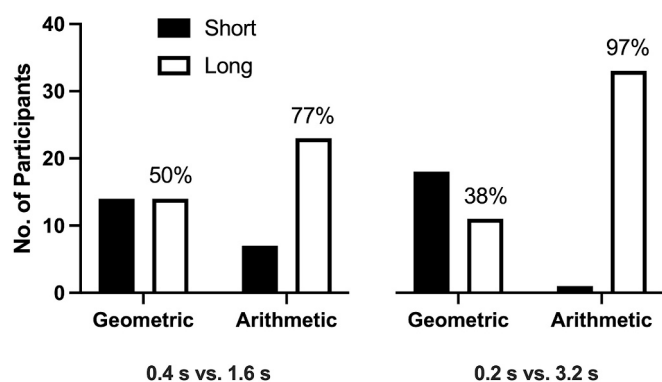
Fig. 3. The number of participants that made a 'short' or 'long' response in the probe test of Experiment 2. The left panel shows the results for the 0.4 s vs. 1.6 s group. The right panel shows the results for the 0.2 s vs. 3.2 s group. The percentages reported above white bars indicates the proportion of participants that responded 'long'.

were compared within the two groups (group 0.4 s vs. 1.6 s and group 0.2 s vs. 3.2 s). The third comparison was of the distribution of responses for participants tested with the arithmetic mean in groups 0.4 s vs. 1.6 s and 0.2 s vs. 3.2 s. The fourth comparison was of the responses to the geometric mean across the two groups. The family-wise error rate for the four comparisons was reduced by adjusting alpha using the Holm-Bonferroni correction. Uncorrected *p*-values are reported as well as the value of alpha of the relevant comparison.

There was a significant association between responses and mean duration for the 0.2 s vs. 3.2 s group, $\chi^2 (63) = 25.98$, $p < 0.001$, alpha = 0.0125, $\varphi = 0.64$. For the 0.4 s vs. 1.6 s group, the association failed to be significant when corrected for multiple comparisons, $\chi^2 (58) = 4.46$, $p = 0.035$, alpha = 0.025 $\varphi = 0.28$. A comparison of the two groups tested with the arithmetic mean showed a significant association between group and response, $\chi^2 (64) = 6.06$, $p = 0.014$, alpha = 0.017, $\varphi = 0.31$. There was no significant association of factors for groups tested with the geometric mean, $\chi^2 (57) < 1$, $p = 0.4$, alpha = 0.05.

Binomial tests showed that for both groups, a significant proportion of participants responded 'long' when presented with the arithmetic mean (largest p-value = 0.005). The proportion of participants that responded 'long' when presented with the geometric mean was not significant for either group (smallest p-value = 0.26).

The analysis of the 0.4 s vs. 1.6 s group, the condition that replicated the design of Experiment 1, failed to be significant when corrected for multiple comparisons. In order to examine the strength of evidence that the results of group 0.4 s vs. 1.6 s replicated the findings of Experiment 1, a replication Bayes factor (BF) was calculated (Ly, Etz, Marsman, & Wagenmakers, 2019) in which the BF for the combined data from group 0.4 s 1.6 s and those of Experiment 1 ($BF_{10} = 272.42$) is expressed as a ratio of the BF for the data of Experiment 1 alone ($BF_{10} = 24.39$). The replication BF_{10} equalled 11.17, indicating that the evidence of replication was 11.17 times greater than the evidence for no replication.

Similar to the findings of Experiment 1, the results of Experiment 2 were not consistent with bisection at the arithmetic mean and were, instead, consistent with bisection at the geometric mean. Both groups showed a preference for making 'long' responses when presented with the arithmetic mean and this preference was significantly greater for the 0.2 s vs. 3.2 s group than 0.4 s vs. 1.6 s group. This is consistent with the hypothesis that bisection occurs at the geometric mean. Because the geometric mean was matched across groups, this resulted in the arithmetic mean of the 0.2 s vs. 3.2 s group being longer than the 0.4 s vs. 1.6 s group and, therefore, increasing the probability that it would be categorised as 'long'.

4. Experiment 3

The purpose of Experiment 3 was to test bisection with a greater stimulus range than compared to Experiments 1 and 2. Similar to Experiment 2, the 0.4 s vs. 1.6 s condition was replicated for a group of participants and was included as a positive control. For another group, the training durations were 0.4 s vs. 6.4 s.

4.1. Method

4.1.1. Participants

113 participants were recruited in the same manner as Experiments 1 and 2. Seventeen were male, 95 were female and 1 participant did not report their sex. Of the 112 participants that reported their age, the mean age was 20 years (range: 18–29).

4.1.2. Apparatus and stimuli

The apparatus and stimuli were the same as Experiment 1 and 2 except that the visual stimulus was a six-pointed purple star.

4.1.3. Procedure

Participants were randomly allocated to either the 0.4 s vs. 1.6 s

condition or the 0.4 s vs. 6.4 s condition. In the probe trial, participants in the 0.4 s vs. 1.6 s condition were randomly allocated either the geometric mean of 0.8 s or the arithmetic mean of 1 s. In the 0.4 s vs. 6.4 s condition, participants were randomly allocated the geometric mean of 1.6 s or the arithmetic mean of 3.4 s. All other details were the same as Experiment 2.

4.2. Results and discussion

Seventeen participants were excluded for failing to perform above 50% during training. The mean performance of the remaining 96 participants was 89% correct, SEM = 1.2.

The distribution of 'short' and 'long' responses to the geometric and arithmetic mean probe durations is shown in Fig. 4. For both the 0.4 s vs. 1.6 s group and the 0.4 s vs. 6.4 s group, participants predominantly responded 'long' when presented with the arithmetic mean. This was not the case for participants presented with the geometric mean of the durations.

Similar to the analyses of Experiment 2, four separate Chi square analyses were conducted, using the Holm-Bonferroni correction, to compare the distribution of 'short' and 'long' responses amongst the groups. For the 0.4 s vs. 1.6 s group, there was a significant association between responses and mean duration, χ^2 (46) = 7.16, $p = 0.008$, alpha = 0.017, $\varphi = 0.39$. The association was also significant for the 0.4 s vs. 6.4 s group, χ^2 (50) = 7.51, $p = 0.006$, alpha = 0.0125, $\varphi = 0.39$. A comparison of the two groups tested with the geometric mean failed to find a significant association between factors, χ^2 (51) = 3.36, $p = 0.07$, alpha = 0.05. A comparison of the two groups tested with the arithmetic mean failed to find a significant association between factors, χ^2 (45) = 3.67, $p = 0.06$, alpha = 0.025.

Binomial tests showed that for both groups a significant proportion of participants responded 'long' when presented with the arithmetic mean (largest p-value = 0.0015). The proportion of participants that responded 'long' when presented with the geometric mean was not significant for the 0.4 s vs. 1.6 s group ($p = 1.0$). The proportion of 'long' responses was significantly greater than short responses for the 0.4 s vs. 6.4 s group ($p = 0.029$).

In order to test the strength of evidence that the results of group 0.4 s vs. 1.6 s replicated the results of Experiment 1 and the similar condition in Experiment 2, a replication BF was calculated. For the combined results of Experiments 1–3, the $BF_{10} = 5156$. For the combined results of Experiments 1 and 2 the $BF_{10} = 24.39$. Therefore, the replication BF_{10} equalled 18.93, indicating that evidence of replication was 18.93 times greater than the evidence for no replication.

The results of Experiment 3 were very similar to Experiment 2 and showed evidence against bisection at the arithmetic mean regardless of

Fig. 4. The number of participants that made a 'short' or 'long' response in the probe test of Experiment 3. The left panel shows the results for the 0.4 s vs. 1.6 s group. The right panel shows the results for the 0.4 s vs. 6.4 s group. The percentages reported above white bars indicates the proportion of participants that responded 'long'.

the stimulus range used. While the results are more consistent with bisection at the geometric mean than the arithmetic mean, there was some evidence for a preference for responding 'long' over 'short' to the geometric mean for the 0.4 s vs. 6.4 s group. This suggests that at some stimulus durations bisection may be below the geometric mean, potentially closer to the harmonic mean. Bisection at the harmonic mean would suggest processing of rate information rather than simply duration. The fact that sub-geometric mean bisection was observed with a long-to-short duration ratio of 16:1 is surprising given that it has previously been observed only with much smaller ratios (Kopec & Brody, 2010). Although, there was some evidence of a preference for 'long' responses, the distribution of responses to the geometric mean for the 0.4 s vs. 6.4 s group, however, did not significantly differ from those of the 0.4 s vs. 1.6 s group. Therefore, the evidence for sub-geometric bisection should be treated with caution.

The finding that participants categorised the arithmetic mean as 'long' to a greater extent than the geometric mean in the 0.4 s versus 1.6 s condition in Experiments 1 and 2 was replicated in the present experiment. Therefore, this finding is robust despite there being only a small difference between the geometric mean of 0.8 s and the arithmetic mean of 1.0 s. In order to determine the effect size for the 0.4 s versus 1.6 s condition, given all the available data, a pooled analysis of the relevant conditions across Experiments 1–3 was conducted. Overall, 79 out of 102 (77%) participants responded 'long' when presented with the arithmetic mean and 52 out of 110 (47%) responded 'long' when presented with the geometric mean. It was found that $\varphi = 0.31$ (odds ratio = 3.8, 95% CI [2.1, 6.9]). A power analysis indicated that samples of $N = 40$ per condition (arithmetic mean and geometric mean) are required for 80% power (alpha = 0.05, two-tailed). A pooled analysis of the conditions that used a larger 16:1 long-to-short duration ratio (0.2 s versus 3.2 s in Experiment 2 and 0.4 s versus 6.4 s in Experiment 3) was also carried out. In contrast to the 0.4 s versus 1.6 s condition, for these conditions there was a larger difference between the geometric and arithmetic means and, therefore, the effect of the probe duration should be larger. It was found that that $\varphi = 0.52$ (odds ratio = 48, 95% CI [7.7, 498] and $N = 15$ per condition is required for 80% power (alpha = 0.05, two-tailed).

5. Experiment 4

The results of Experiments 1–3 provide evidence against temporal bisection at the arithmetic mean and are, instead, consistent with bisection at the geometric mean. This is in contrast to the previous research that supported bisection at the arithmetic mean but assessed bisection with a distribution of probe durations (e.g., Wearden et al., 1997; Wearden & Ferrara, 1995, 1996). Other than the difference in the nature of the probe test, there are potential differences between the methods of Experiments 1–3 and those of other studies that have produced contradictory results. While it is not clear why it would lead to the pattern of the results, the most obvious methodological difference is that the current experiments were conducted online rather than in a controlled, laboratory environment. The purpose of Experiment 4 was to verify that the online procedure used in Experiments 1–3 would replicate bisection at the arithmetic mean when tested with linearly spaced probe durations. In addition, another group of participants were tested with logarithmically spaced probe durations in order to test whether the procedure would replicate the spacing effect such that logarithmically spaced probe durations would lead to shorter bisection points than linearly spaced probe durations. Replication of the spacing effect would suggest that the difference between the results of Experiments 1–3 and previous research showing bisection at the arithmetic mean likely reflects the difference in the probe test procedure rather than other differences in methods.
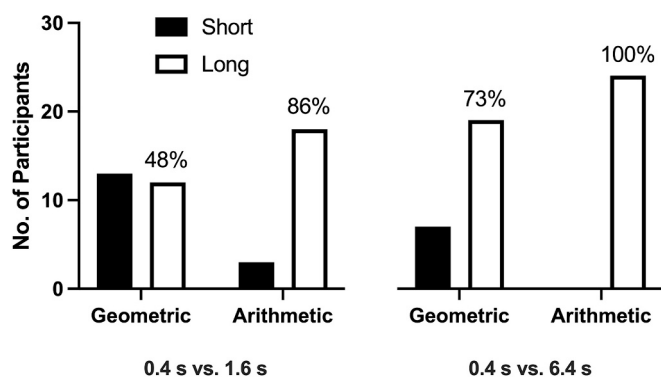
5.1. Methods

5.1.1. Participants

Thirty-seven participants were recruited in the same manner as Experiments 1–3. Four were male, 31 were female and 2 participants did not report their sex. Of the 35 participants that reported their age, the mean age was 20 years (range: 17–23).

5.1.2. Apparatus and stimuli

The apparatus and stimuli were the same as Experiment 1 and 2.

5.1.3. Procedure

Participants were randomly allocated to either the logarithmic or linear conditions. The procedure for training was identical to Experiments 1-3. In the test phase, participants continued to receive training trials with the 0.4 s and 1.6 s durations and received feedback. Interspersed with the training trials were probe trials that consisted of intermediate durations. For the logarithmic spacing group, these probe durations were 0.5 s, 0.6 s, 0.8 s, 1.0 s and 1.3 s. For the linear spacing group, the probe durations were 0.6 s, 0.8 s, 1.0 s, 1.2 s and 1.4 s. No feedback was given on probe duration trials. The test phase consisted of 70 trials: 10 trials of each of the two training durations and the five probe durations. Trials were presented in a random order except that each block of seven trials contained one of each duration and the first two trials of the block were the training durations. All other details were the same as Experiment 1.

5.2. Results and discussion

Six participants were excluded for failing to perform above 50% during training. The mean performance of the remaining 31 participants was 93% correct, SEM = 1.7.

The proportion of 'long' responses for the logarithmic ($N = 12$) and linear ($N = 19$) spacing groups across probe durations is shown in Fig. 5, left panel. The proportion of 'long' responses increased as a function of time for both groups, but the increase was earlier for the logarithmic spacing group compared to the linear spacing group. The bisection point for each participant was determined by fitting a slope to the steepest point of increase in the proportion of 'long' responses over the probe durations (see Maricq, Roberts, & Church, 1981; Wearden & Ferrara, 1996). The steepest point was determined by fitting slopes for the first four shortest probe durations (i.e., 0.4 s, 0.5 s, 0.6 s, 0.8 s for the logarithmic group and 0.4 s, 0.6 s, 0.8 s, 1.0 s for the linear group) and then for the second to fifth shortest durations and so on. The median bisection point for each group is shown in Fig. 5, right panel. The bisection points for the logarithmic group were significantly shorter than those of the linear group, $t(29) = 3.1$, $p = 0.004$, Cohen's d = 1.14, 95% CI = 0.35, 1.91. One-sample t-tests for each group comparing bisection points to
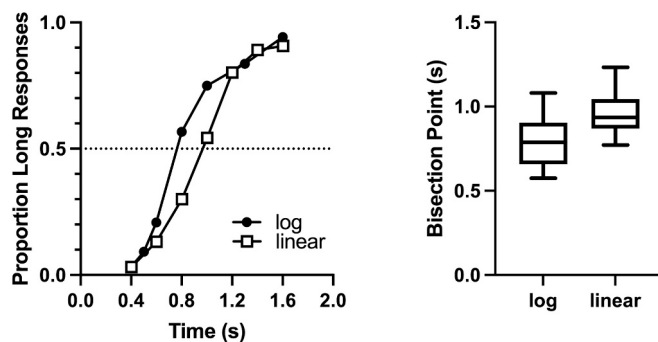
the arithmetic mean showed that whereas the logarithmic group had bisection points significantly below the arithmetic mean, $t(11) = 4.31$, $p = 0.001$, Cohen's d = 1.25, 95% CI = 0.47, 1.99, the linear group did not, $t(18) = 1.25$, $p = 0.11$.

The results replicated the finding that bisection occurs close to the arithmetic mean in humans with linearly spaced probe durations and that logarithmically spaced probe durations result in a reduction of the bisection point. While the result of the logarithmically spaced condition is consistent with the findings of Experiments 1–3, the result of the linearly spaced condition is not.

6. General discussion

Across Experiments 1–3, it was consistently found that the arithmetic mean of the training durations was preferentially categorised as long. This provides evidence against the hypothesis that the perceived midpoint of two durations is the arithmetic mean in humans. The effect was repeatedly found with training durations of 0.4 s and 1.6 s. It was also found with a much larger 16:1 ratio between the durations in Experiments 2 and 3, regardless of whether the durations were 0.2 s vs. 3.2 s or 0.4 s vs. 6.4 s. The procedures used in Experiments 1–3 avoided the confound of probe duration distribution on test performance. Therefore, the results suggest that bisection at the arithmetic mean may be a performance effect reflecting a bias to distribute responding equally over the two response options. The results of Experiments 1–3 were not an artefact of the procedure because Experiment 4, in which the same training procedure was used, replicated the probe duration spacing effect and demonstrated bisection close to the arithmetic mean with linearly spaced probe durations.

Before considering the implications of the results for hypotheses for the bisection point, it is important to, first, consider potential alternative interpretations of the data. The use of the temporal bisection procedure rests on the assumption that subjects compare an experienced duration, in the probe phase, to the memory of both the short and long durations. There is the potential, however, that successful discrimination of short and long durations may be achieved by memory of the short duration alone. Thus, it is possible that participants compare experienced durations to the short duration and once the experienced duration has passed a criterion (e.g., a multiple of the short duration) then the duration is categorised as being long. Such a decision process has been used successfully to account for the results of temporal generalisation studies in animals and humans in which the comparison of durations is limited to one trained duration (Church & Gibbon, 1982; Wearden, 1992). It is unlikely, however, that this decision process is used in tests of temporal bisection, because it is typically found that the range of short and long durations affects the bisection point (Wearden & Ferrara, 1996) and, thus, manipulation of either the short or long durations while holding the other constant affects bisection (Kopec & Brody, 2010). Therefore, the bisection point reflects sensitivity to both the short and long durations. In the present experiments, there was no significant evidence that the use of this rule led to the observed results. In Experiment 2, the two groups were trained with different short and long durations (either 0.4 s versus 1.6 s or 0.2 s versus 3.2 s). The geometric mean of the intervals, however, was the same for both groups. If responding was based solely on comparison with the short duration, then it would be anticipated that the group trained with 0.2 s short duration would show a greater proportion of 'long' responses than the group trained with the 0.4 s short duration when tested with the geometric mean. This was not the case and there was no significant difference between the groups. Similarly, in Experiment 3, because both groups were trained with a short duration of 0.4 s, a response criterion based solely on the short duration would result in the group trained with a 6.4 s long duration making a greater proportion of 'long' responses than the group trained with the 1.6 s long duration. This was not the case and there was no significant effect of long duration for either the geometric or arithmetic mean probe conditions.

Fig. 5. Test performance in Experiment 4. The left panel shows the mean proportion of 'long' responses as a function of time for the logarithmic (log) and linear groups. The right panel shows the median, inter-quartile range and minimum and maximum bisection points for the two groups.

The results suggest that the perceived midpoint of temporal durations is closer to the geometric mean than it is to the arithmetic mean. This is consistent with findings in non-human animals that show bisection close to the geometric mean, e.g., in rats (Church & Deluty, 1977) and pigeons (Platt & Davis, 1983; Stubbs, 1976). Bisection at the geometric mean may be evidence of logarithmic perception of time (Church & Deluty, 1977; Ren, Allenmark, Muller, & Shi, 2020). Logarithmic encoding has been suggested to account for discrimination effects in other stimulus domains such as in vision (Mackay, 1963) and fits with Weber's law that the just noticeable difference between stimuli reflects the relative rather than absolute properties of the stimuli. Logarithmic perception of time results in bisection at the geometric mean if people compare the differences between durations. Thus, the point that is half-way between two values on a logarithmic scale is the geometric mean of the absolute values. Bisection at the geometric mean, however, may also occur as a consequence of a linear relationship between physical and perceived time. Gibbon (1977) proposed that time is encoded linearly but animals compare the ratio of durations. This results in bisection between two durations occurring at the duration that is a ratio of the short duration that is equal to the square root of the ratio of the long and short durations. The present results are not able to differentiate between the two accounts. Gibbon and Church (1981), however, favoured an account of linear encoding based on the simultaneous comparison of an interval with a partially elapsed interval. In their study, rats were trained to respond to one lever that was reinforced after 60 s and another lever that was reinforced after 30 s. On probe trials, the lever that was reinforced after 30 s was presented half-way through the presentation of the lever that was reinforced after 60 s. Rats showed similar levels of responding to the two levers indicating that the expected time to reinforcement was the same for both response options. This would not be the case if time was perceived logarithmically, as this would result in the expected time to reinforcement being shorter for the lever reinforced after 60 s rather than the lever reinforced after 30 s. A similar study in humans has also favoured a linear encoding account (Wearden, 2002). These results are potentially in contrast to the findings of ratio setting experiments in which participants were asked to indicate the duration that is a particular proportion of another duration, e.g., the duration that is either half or double that of 1 s (Allan, 1978; Eisler, 1976). While these studies were consistent with logarithmic perception of time, this property of responding may be caused by biases in the subjective proportions used by participants rather than the relation between subjective and physical time (Allan, 1978). Regardless of the mechanisms by which bisection close to the geometric mean occurs, the results demonstrate that humans and animals similarly bisect below the arithmetic mean demonstrating that, at the least, time categorisation does not reflect the combination of linear encoding and a comparison of the difference of the durations.

A number of approaches for accounting for the difference between the animal and human studies have been proposed. These have tended to assume that, for humans, the perceived midpoint is the arithmetic mean, but that under certain conditions, bisection may occur below the arithmetic mean due to potential response biases (e.g., Wearden, 1991). The current results do not fit with such models because the results show that the opposite is true: the perceived midpoint is closer to the geometric mean than arithmetic mean, but bisection may occur close to the arithmetic mean due to a response bias. This difference is clear when considering a model proposed by Kopec and Brody (2010) that accounts for data from human temporal bisection across 18 papers. The results do not fit the model because it assumes that time perception is linear and, all other factors being equal, that bisection will occur at the arithmetic mean. The model successfully accounts for other data suggesting that, on the whole, bisection is close to the arithmetic mean but often sub-arithmetic and as the ratio of the durations decreases bisection will become closer to the geometric mean and eventually sub-geometric. Sub-arithmetic bisection is proposed to occur as a consequence of a series of processes. First, there is a greater probability of short durations being classed as neither the 'short' or 'long' durations. This occurs because the distribution of the probability of classifying a duration as the long duration is greater than the distribution of the short duration based on the assumption that variance scales with the remembered duration (Gibbon, 1977). Therefore, intermediate durations closer to the trained short duration are more likely to be determined as being ambiguous and, subsequently, responding is determined by how similar the probe duration is to the remembered short and long duration. Second, ambiguous, intermediate durations are subject to a response bias based on the gambler's fallacy such that the tendency to make a particular response decreases as a function of the frequency with which it has already been made. Because more short durations than long durations are classified as ambiguous, initially, in the test phase, responses to durations classified as ambiguous will be predominantly 'short'. Subsequently, the tendency to respond 'short' decreases and more intermediate durations are classified as long such that the bisection point is reduced below the arithmetic mean. When the long-to-short ratio is small, the difference between the arithmetic mean and geometric mean is small. Under these circumstances, the response bias can lead to sub-geometric bisection. Thus, while the model predicts sub-arithmetic bisection, it assumes that it is a consequence only of a response bias. Consequently, the theory makes the prediction that on the first trial of the probe test phase there is no bias such that the probability of responding 'short' or 'long' to the arithmetic mean will be equal.

Another way in which the results do not fit the Kopec and Brody (2010) model is that it attempts to account for the observation that the bisection point moves closer to the arithmetic mean as the ratio of long-to-short durations gets larger in standard tests of bisection that use a distribution of probe durations. In Experiment 2, the opposite effect was observed. Participants that were trained with 0.2 s and 3.2 s (long-to-short ratio of 16) were more likely to respond 'long' than participants trained with 0.4 s and 1.6 s (long-to-short ratio of 4) when presented with the arithmetic mean. Because both pairs of short and long durations share the same geometric mean of 0.8 s, it would be expected, if bisection occurs close to the geometric mean, that the arithmetic mean of 0.2 s and 3.2 s (1.7 s) would lead to a more extreme rating of 'long' than the arithmetic mean of 0.4 s and 1.6 s (1.0 s). Therefore, the results are unlikely to fit with any model that assumes that bisection moves away from the geometric mean towards the arithmetic mean as the ratio of the durations increases.

A recent account of variation in human temporal bisection is that it reflects sensitivity to the ensemble statistics of the experienced durations (Zhu et al., 2021). People represent the mean and distribution of the experience durations. Probe durations are compared to these statistics rather than directly with the trained short and long durations. This account, however, made no specific assumptions about whether the mean of the distribution is arithmetic or geometric. Zhu et al. (2021) concluded that the assumption of the geometric mean or arithmetic made little difference in the ability of the model to account for the data. While this may be true across a distribution of probe duration trials, the assumptions about the nature of the represented mean of the distribution will have a large effect on the first probe trial when the mean is based on only the trained durations, the extreme values of the sample distribution. Based on the current experiments, it must be concluded that the ensemble mean is not the arithmetic mean.

Differences in human and non-human animal bisection have also been suggested to reflect differences in the degree of temporal discounting (Kopec & Brody, 2018). The assumption is that the underlying representation of time across species is linear, but animals show greater temporal discounting of rewards resulting in sub-arithmetic temporal bisection. The current results suggest that humans, like animals, show sub-arithmetic bisection when response biases are avoided in the test phase. Therefore, the differences between species can be reconciled without appealing to differences in sensitivity to temporal discounting. Bisection above the geometric mean is the result of a bias to distribute responding across the two response options. This may be due to a

central-tendency effect that reflects sampling of the probe duration distribution (Lejeune & Wearden, 2009). Alternatively, it may reflect learning of the distribution of responses to the two response options during the learning phase that affects performance in the test phase when the distribution of probe durations results in a change in the proportion of 'short' and 'long' responses over trials (Cambraia, Vasconcelos, Jozefowiez, & Machado, 2021; Jozefowiez, Polack, Machado, & Miller, 2014).

While Experiments 1–3 removed the influence of the distribution of probe trial durations, it is likely that participants recalled the trained durations with a certain amount variance in the estimates of time such that probe durations were always compared to a distribution of durations that was greater than just the trained durations. The results of the experiments indicate that, at the very least, a difference comparison was not used in conjunction with internal representations of time linearly distributed between short and long durations. It is also possible that people form an internal reference duration when asked to make temporal categorisation judgements (Bausenhart, Bratzke, & Ulrich, 2016), reflecting a internal estimate of the average of the durations that have been experienced. This would suggest the bisection point is not simply the point of subjective equality, but the point that matches the internal representation of the average. The results of Experiments 1–3 suggest that the internal reference duration is not the arithmetic mean, which, once again, rules out the possibility that temporal discrimination reflects comparison of the differences of linearly represented durations.

Although the temporal bisection procedure may be used to assess time perception, the nature of the procedure means that it relies on memory and requires the comparison of elapsed time and stored mnemonic representations of time. This issue has been recognised by theories of time perception that identify multiple cognitive factors that contribute to temporal perception (e.g., Gibbon, 1977; Killeen & Grondin, 2022; Staddon & Higa, 1999). Furthermore, it is a challenge, for the investigation of impaired timing, to identify the precise cause of the impairment (Allman & Meck, 2012) considering it may arise from a number of different cognitive processes. While the current results may be limited to falsifying the hypothesis that the bisection point is the arithmetic mean, they reconcile the human temporal bisection literature with other literature consistent with scalar properties of timing (Haigh et al., 2021; Rakitin et al., 1998) suggesting that memory for time and the decision processes involved in temporal discrimination result in a seemingly non-linear relation between subjective and physical time.

CRediT authorship contribution statement

David J. Sanderson: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Resources, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

None.

Data availability

The data are are available as Supplementary data.

Acknowledgements

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2024.105770.

References

Allan, L. G. (1978). Comments on current ratio-setting models for time perception. *Perception & Psychophysics, 24*(5), 444–450. https://doi.org/10.3758/bf03199742

Allan, L. G. (2002). The location and interpretation of the bisection point. *The Quarterly Journal of Experimental Psychology. B, 55*(1), 43–60. https://doi.org/10.1080/02724990143000162

Allman, M. J., & Meck, W. H. (2012). Pathophysiological distortions in time perception and timed performance. *Brain, 135*(Pt 3), 656–677. https://doi.org/10.1093/brain/awr210

Bausenhart, K. M., Bratzke, D., & Ulrich, R. (2016). Formation and representation of temporal reference information. *Current Opinion in Behavioral Sciences, 8*, 46–52.

Brown, G. D., McCormack, T., Smith, M., & Stewart, N. (2005). Identification and bisection of temporal durations and tone frequencies: Common models for temporal and nontemporal stimuli. *Journal of Experimental Psychology. Human Perception and Performance, 31*(5), 919–938. https://doi.org/10.1037/0096-1523.31.5.919

Cambraia, R., Vasconcelos, M., Jozefowiez, J., & Machado, A. (2021). Base rates bias performance in a temporal bisection task. *The Journal of Experimental Psychology: Animal Learning and Cognition, 47*(2), 163–182. https://doi.org/10.1037/xan0000284

Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology. Animal Behavior Processes, 3*(3), 216–228. https://doi.org/10.1037//0097-7403.3.3.216

Church, R. M., & Gibbon, J. (1982). Temporal generalization. Journal of Experimental Psychology. Animal Behavior Processes, 8(2), 165–186. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7069377.

Eisler, H. (1976). Experiments on subjective duration 1968-1975: A collection of power function exponents. Psychological Bulletin, 83(6), 1154–1171. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/996212.

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

Gibbon, J. (1981). On the form and location of the psychometric bisection function for time. *Journal of Mathematical Psychology, 24*, 58–87.

Gibbon, J., & Church, R. M. (1981). Time left: Linear versus logarithmic subjective time. Journal of Experimental Psychology. Animal Behavior Processes, 7(2), 87–107. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7241054.

Haigh, A., Apthorp, D., & Bizo, L. A. (2021). The role of Weber's law in human time perception. *Attention, Perception, & Psychophysics, 83*(1), 435–447. https://doi.org/10.3758/s13414-020-02128-6

Hecht, S. (1934). A theoretical basis for intensity discrimination in vision. *Proceedings of the National Academy of Sciences of the United States of America, 20*(12), 644–655. https://doi.org/10.1073/pnas.20.12.644

JASP-Team (2024). JASP (Version 0.17.2.1). Retrieved from https://jasp-stats.org/.

Jozefowiez, J., Polack, C. W., Machado, A., & Miller, R. R. (2014). Trial frequency effects in human temporal bisection: Implications for theories of timing. *Behavioural Processes, 101*, 81–88. https://doi.org/10.1016/j.beproc.2013.07.023

Killeen, P. R., & Grondin, S. (2022). A trace theory of time perception. *Psychological Review, 129*(4), 603–639. https://doi.org/10.1037/rev0000308

Kopec, C. D., & Brody, C. D. (2010). Human performance on the temporal bisection task. *Brain and Cognition, 74*(3), 262–272. https://doi.org/10.1016/j.bandc.2010.08.006

Kopec, C. D., & Brody, C. D. (2018). Controlling for temporal discounting shifts rats from geometric to human-like arithmetic bisection. *bioRxiv*. https://doi.org/10.1101/245092

Lejeune, H., & Wearden, J. H. (2009). Vierordt's the experimental study of the time sense (1868) and its legacy. *European Journal of Cognitive Psychology, 21*(6), 941–960. https://doi.org/10.1080/09541440802453006

Ly, A., Etz, A., Marsman, M., & Wagenmakers, E. J. (2019). Replication Bayes factors from evidence updating. *Behavior Research Methods, 51*(6), 2498–2508. https://doi.org/10.3758/s13428-018-1092-x

Mackay, D. M. (1963). Psychophysics of perceived intensity: A theoretical basis for Fechner's and Stevens' Laws. *Science, 139*, 1213–1216.

Maricq, A. V., Roberts, S., & Church, R. M. (1981). Methamphetamine and time estimation. *Journal of Experimental Psychology. Animal Behavior Processes, 7*(1), 18–30. https://doi.org/10.1037//0097-7403.7.1.18

Moore, B. C., & Raab, D. H. (1975). Intensity discrimination for noise bursts in the presence of a continuous, bandstop background: Effects of level, width of the bandstop, and duration. *The Journal of the Acoustical Society of America, 57*(2), 400–405. https://doi.org/10.1121/1.380455

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Hochenberger, R., Sogo, H., & Lindelov, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Penney, T. B., Brown, G. D., & Wong, J. K. (2014). Stimulus spacing effects in duration perception are larger for auditory stimuli: Data and a model. *Acta Psychologica, 147*, 97–104. https://doi.org/10.1016/j.actpsy.2013.07.017

Penney, T. B., & Cheng, X. (2018). Duration bisection: A User's guide. In A. Vatakis, F. Balci, M. Di Luca, & A. Correa (Eds.), *Timing and time perception: Procedures, measures, & applications* (pp. 98–127). Brill.

Platt, J. R., & Davis, E. R. (1983). Bisection of temporal intervals by pigeons. Journal of Experimental Psychology. Animal Behavior Processes, 9(2), 160–170. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/6842135.

Rakitin, B. C., Gibbon, J., Penney, T. B., Malapani, C., Hinton, S. C., & Meck, W. H. (1998). Scalar expectancy theory and peak-interval timing in humans. *Journal of Experimental Psychology. Animal Behavior Processes, 24*(1), 15–33. https://doi.org/10.1037//0097-7403.24.1.15

Raslear, T. G. (1983). A test of the Pfanzagl bisection model in rats. Journal of Experimental Psychology. Animal Behavior Processes, 9(1), 49–62. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/6842132.

Raslear, T. G. (1985). Perceptual bias and response bias in temporal bisection. *Perception & Psychophysics, 38*(3), 261–268. https://doi.org/10.3758/bf03207153

Ren, Y., Allenmark, F., Muller, H. J., & Shi, Z. (2020). Logarithmic encoding of ensemble time intervals. *Scientific Reports, 10*(1), 18174. https://doi.org/10.1038/s41598-020-75191-6

Staddon, J. E., & Higa, J. J. (1999). Time and memory: Towards a pacemaker-free theory of interval timing. *Journal of the Experimental Analysis of Behavior, 71*(2), 215–251. https://doi.org/10.1901/jeab.1999.71-215

Stubbs, D. A. (1976). Scaling of stimulus duration by pigeons. *Journal of the Experimental Analysis of Behavior, 26*(1), 15–25. https://doi.org/10.1901/jeab.1976.26-15

Wearden, J. H. (1991). Human performance on an analogue of an interval bisection task. The Quarterly Journal of Experimental Psychology. B, 43(1), 59–81. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/2017575.

Wearden, J. H. (1992). Temporal generalization in humans. *Journal of Experimental Psychology: Animal Behavior Processes, 18*(2), 134–144.

Wearden, J. H. (2002). Traveling in time: A time-left analogue for humans. Journal of Experimental Psychology. Animal Behavior Processes, 28(2), 200–208. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/11987876.

Wearden, J. H., & Ferrara, A. (1995). Stimulus spacing effects in temporal bisection by humans. The Quarterly Journal of Experimental Psychology. B, 48(4), 289–310. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/8532899.

Wearden, J. H., & Ferrara, A. (1996). Stimulus range effects in temporal bisection by humans. *The Quarterly Journal of Experimental Psychology. B, 49*(1), 24–44. https://doi.org/10.1080/713932615

Wearden, J. H., Rogers, P., & Thomas, R. (1997). Temporal bisection in humans with longer stimulus durations. *The Quarterly Journal of Experimental Psychology. B, 50*(1), 79–94. https://doi.org/10.1080/027249997393655

Zhu, X., Baykan, C., Muller, H. J., & Shi, Z. (2021). Temporal bisection is influenced by ensemble statistics of the stimulus set. *Attention, Perception, & Psychophysics, 83*(3), 1201–1214. https://doi.org/10.3758/s13414-020-02202-z