Published in final edited form as:

Brain Cogn. 2010 December ; 74(3): 262–272. doi:10.1016/j.bandc.2010.08.006.

Human Performance on the Temporal Bisection Task

Charles D. Kopec and Carlos D. Brody

HHMI / Princeton University, Princeton Neuroscience Institute, Princeton, NJ 08544

Abstract

The perception and processing of temporal information are tasks the brain must continuously perform. These include measuring the duration of stimuli, storing duration information in memory, recalling such memories, and comparing two durations. How the brain accomplishes these tasks however, is still open for debate. The temporal bisection task, which requires subjects to compare temporal stimuli to durations held in memory, is perfectly suited to address these questions. Here we perform a meta-analysis of human performance on the temporal bisection task collected from 148 experiments spread across 18 independent studies. With this expanded data set we are able to show that human performance on this task contains a number of significant peculiarities, which in total no single model yet proposed has been able to explain. Here we present a simple 2-step decision model that is capable of explaining all the idiosyncrasies seen in the data.

Keywords

Temporal Bisection Task; Human; Timing; Perception; Decision Making; Meta-Analysis; Model

INTRODUCTION

Time is the most fundamental component of the world we live in, yet its perception is also the least understood. Even simple everyday tasks such as playing a game of catch, which requires one to generate complex motor outputs (throwing) and predict the future location of a moving object (catching) are intertwined with time. The generation and comprehension of speech requires estimates of time that span from hundreds of milliseconds, for the comprehension of syllables, to tens of minutes, for the comprehension of a conversation. Time pervades all aspects of our lives, and is therefore something our brains must continuously deal with. Yet amazingly, a problem so fundamental, remains largely a mystery.

Over the past half century our understanding of the neurobiological nature of various cognitive functions has expanded enormously. Systems such as vision, audition, and somatosensory perception have been unraveled by tracing the neural circuitry and activity from their primary sensory organs to their specialized regions of neocortex. The study of time however has lagged behind. Without an obvious temporal sensory organ to start from, such neural tracing experiments could not be conducted. In their absence we are left with a

© 2010 Elsevier Inc. All rights reserved.

Corresponding Author: Charles D. Kopec, HHMI / Princeton University, Lewis Thomas Lab, Princeton, NJ 08544, ckopec@princeton.edu, Phone: 609-258-7316, Fax: 609-258-1028.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

wealth of psychological studies conducted over the past two decades that through their ingenious design have attempted to unlock the mystery of time in the brain.

The temporal bisection task, originally used by Russell Church and Marvin Deluty in 1977 to study temporal discrimination in rats (Church & Deluty, 1977), was first applied to humans in two separate pioneering works by John Wearden (Wearden, 1991) and Lorraine Allan and John Gibbon (L.G. Allan & Gibbon, 1991) in 1991 (though a simpler version was used on humans as early as 1968 (Bovet, 1968)). In the task, subjects are required to compare temporal stimuli to two reference stimuli, “long” and “short”, held in memory. The stimuli themselves are generally either a tone or a light presented for some length of time. Generally, subjects are first pretrained on the reference stimuli, after which intermediate probe stimuli are introduced (Figure 1A). In other versions the reference stimuli are never explicitly identified (Droit-Volet & Rattat, 2007; Wearden & Ferrara, 1995), but still are the shortest and longest stimuli presented. Upon being presented with an intermediate probe stimulus, the subject must indicate which reference stimulus they believe it is more similar to. If the reference stimuli were never specifically identified, the subject must simply classify the duration as “short” or “long”. This is a subjective decision so no feedback is given as there is no correct or incorrect answer.

This task is ideal for studying the perception and processing of temporal information because it requires subjects to perform a number of time-dependent mental operations. First the “short” and “long” reference durations (denoted here as TS and TL, respectively) must be learned and stored in memory. Second, the length of the probe duration must be measured. Third, the values of the “short” and “long” reference durations must be retrieved from memory. Finally, the probe duration must be compared to the reference durations and a decision reached.

From the data collected during the task, experimenters are able to construct a psychometric curve plotting the duration of the stimuli (probes and reference stimuli) versus the subject’s probability of responding “long” (Figure 1B). These functions show a monotonic increase with duration, meaning subjects almost never respond “long” to the shortest duration (namely, the “short” reference duration), and almost always respond “long” to the longest duration (namely, the “long” reference duration). At some intermediate duration, the subject’s performance crosses 0.5 on the y-axis. It is this duration, referred to as the bisection point or point of indifference (L.G. Allan & Gibbon, 1991; Church & Deluty, 1977; Gibbon, 1981; Siegel & Church, 1984; Wearden, 1991), that they are equally likely to call “long” or “short”. This single, seemingly trivial, point actually offers significant insight to how time is represented and processed in the brain, because at this duration the decision process used to compare temporal stimuli to temporal values stored in memory must be equal for both options.

Starting in the late 1970s and throughout the 80s the Temporal Bisection Task was exclusively performed with non-human subjects (Church & Deluty, 1977; Gibbon, 1981; Meck, 1983; Raslear, 1983; Siegel, 1986; Siegel & Church, 1984). Many of the features present in human performance discussed here have parallels with earlier discoveries in the animal literature. However, a number of key differences between human and non-human performance on the temporal bisection task have also emerged (Church & Deluty, 1977; Gibbon, 1981; Siegel, 1986). It is for this reason that we restrict our analysis and modeling to human data only.

Over the past two decades many insightful studies performing the temporal bisection task on human subjects have been published (L. G. Allan, 2002; L. G. Allan & Gerhardt, 2001; L.G. Allan & Gibbon, 1991; Bovet, 1968; Droit-Volet, 2003; Droit-Volet, Meck, & Penney,

2007; Droit-Volet & Rattat, 2007; Droit-Volet, Tourret, & Wearden, 2004; Droit-Volet & Wearden, 2001; Lieving, Lane, Cherek, & Tcheremissine, 2006; Nichelli, Alway, & Grafman, 1996; Ortega & Lopez, 2008; Penney, Gibbon, & Meck, 2000; Penney, Meck, Roberts, Gibbon, & Erlenmeyer-Kimling, 2005; Wearden, 1991; Wearden & Ferrara, 1995, 1996; Wearden, Rogers, & Thomas, 1997), highlighting both the importance of this task in the field of human time perception and the complexity of human performance on it. Together these studies span a wide range of absolute and relative durations for the reference durations. Each experiment contains one of two different rules governing the spacing of probe durations (i.e. uniformly spaced on a linear or logarithmic axis), and the experiments cover a large range of different subject ages. Here we combine the data into a single data set. With such a large data set we are able to review and test previously described correlations as well as uncover previously unrecognized statistically significant peculiarities in human performance on this task, all of which no single model yet proposed can explain. We then present a novel 2-step decision model that is capable of reproducing each of these peculiarities. Finally, we hope to reveal areas in parameter space that are less heavily studied or even lacking all together, thus motivating and focusing future studies so as to help fill in the data set. From this, more accurate and realistic models of human performance can be produced in the future.

RESULTS

Compiled data

Over the past two decades a number of studies have performed the temporal bisection task on human subjects (L. G. Allan, 2002; L. G. Allan & Gerhardt, 2001; L.G. Allan & Gibbon, 1991; Bovet, 1968; Droit-Volet, 2003; Droit-Volet et al., 2007; Droit-Volet & Rattat, 2007; Droit-Volet et al., 2004; Droit-Volet & Wearden, 2001; Lieving et al., 2006; Nichelli et al., 1996; Ortega & Lopez, 2008; Penney et al., 2000; Penney et al., 2005; Wearden, 1991; Wearden & Ferrara, 1995, 1996; Wearden et al., 1997). Here we have compiled the results from those studies into a single data set (downloaded at http://brodylab.princeton.edu/wiki/images/2/28/Human_Temporal_Bisection_Task_Data.zip). This data set is drawn from 18 independent studies, which conducted a collective 148 experiments involving roughly 1020 subjects performing a total of over 302,000 individual trials of the temporal bisection task. The short reference durations range from 50ms to 8 seconds and the long reference durations range from 200ms to 32 seconds. The spread, or ratio of the reference durations ("Long" / "Short"), ranges from 1.2 to 19. In most studies the durations of the probe trials are either spaced linearly or logarithmically between the reference durations. Stimuli were either a visual cue presented on a computer screen for some length of time, or an auditory tone. Subject's age is classified as 3, 5, 8, or adult.

Throughout the studies compiled here, a number of interesting phenomena were reported in the subjects' performance. With this expanded data set we will examine each of these phenomena and determine whether they are statistically significant across studies. The five studies that examined bisection in children (Droit-Volet, 2003; Droit-Volet et al., 2007; Droit-Volet & Rattat, 2007; Droit-Volet et al., 2004; Droit-Volet & Wearden, 2001) found that performance of 3 and 5-year-olds was significantly different from 8-year-olds and adults. Therefore, all analyses conducted here will exclude children under 8 years of age. We hope that future studies on children will be performed with an expanded range of reference duration spreads (the majority performed to date used only a 4-fold spread) and probe spacing so that the analysis performed below can be extended across development.

Geometric to Arithmetic Transition

The curious fact that the ratio of the “Long” and “Short” reference durations, the spread, affects a subject’s bisection point was noticed in a number of studies (L. G. Allan, 2002; Wearden & Ferrara, 1996). Specifically, when the spread is large ($TL / TS \geq 4$) subjects tend to bisect (B) near the arithmetic mean of the reference durations ($B = (TL + TS) / 2$). However, when the spread is small ($TL / TS \leq 2$) subjects tend to bisect near the geometric mean of the reference durations ($B = \sqrt{TL \times TS}$). To determine whether this observation is significant across studies we analyzed the 88 experiments that used linear spaced probes on subjects 8 years of age and older (L. G. Allan, 2002; L. G. Allan & Gerhardt, 2001; L.G. Allan & Gibbon, 1991; Droit-Volet, 2003; Droit-Volet & Rattat, 2007; Droit-Volet et al., 2004; Droit-Volet & Wearden, 2001; Nichelli et al., 1996; Ortega & Lopez, 2008; Wearden, 1991; Wearden & Ferrara, 1995, 1996; Wearden et al., 1997). For each experiment we computed the distance between the average bisection point and the geometric and arithmetic means of the reference durations. In figure 2A we plot the log ratio of these distances [$\log(\text{abs}(B - GM) / \text{abs}(B - AM))$] as a function of the log spread of the reference durations. Values below 0 indicate bisection was closer to the geometric mean while values above 0 indicate bisection closer to the arithmetic mean. This data shows a significant positive correlation ($r = 0.58$, $p < 0.01$) confirming that the relative spacing of the reference durations affects the bisection point. Also, no discontinuous jumps are seen in the data (the data is not better fit by two lines as opposed to one, $p = 0.44$, and the data is not better fit by a sigmoid as compared to a line, $p = 0.44$), indicating a smooth transition from arithmetic-like bisection to geometric-like bisection as the reference duration spread is decreased.

Sub-geometric bisection

Further examination of the data reveals that bisection actually becomes sub-geometric for spreads < 1.8 (19 of 21 experiments, $p = 0.02$, Figure 2B insert). In figure 2B we plot the bisection point, normalized by the geometric mean as a function of log spread. Values below 1 indicate sub-geometric bisection. Again, the data show a smooth transition of bisection point as a function of spread ($r = -0.88$, $p < 0.01$).

Sub-arithmetic bisection

The observation that a subject’s bisection point transitions from arithmetic to geometric to sub-geometric as a function of spread may lead one to suggest that subjects utilize different strategies to perform the temporal bisection task depending on the spread of the reference durations. However, if we plot bisection normalized by the arithmetic mean as a function of log spread (Figure 2C), it becomes obvious that across the entire range of spreads examined here bisection can be simply described as being sub-arithmetic.

In fact, the arithmetic to geometric to sub-geometric transition in bisection may actually be a mathematical artifact of sub-arithmetic bisection. When the spread between two reference durations is large ($TS = 1$ and $TL = 9$), the arithmetic and geometric means are far apart ($GM = 3$ and $AM = 5$). If bisection occurred at 90% the arithmetic mean ($B = 4.5$) it would appear closer to the AM than the GM. However, if the spread was reduced ($TS = 1$ and $TL = 2.5$) the arithmetic and geometric means come closer together ($GM = 1.58$ and $AM = 1.75$). Now bisection at 90% the arithmetic mean ($B = 1.57$) will appear geometric. Finally, reducing the spread even further ($TS = 1$ and $TL = 1.5$, $GM = 1.22$ and $AM = 1.25$) will produce bisection ($B = 1.12$) that appears sub geometric.

Analysis of the data reveals that bisection becomes slightly more sub-arithmetic as a function of spread (Figure 2C, $r = -.20$, $p = 0.014$ that the data does not have a negative correlation coefficient). A similar trend has been reported from non-human subjects

performing the temporal bisection task (Killeen, Fetterman, & Bizo, 1997; Siegel, 1986; Siegel & Church, 1984). Sub-arithmetic bisection may be indicative of an inherent propensity to respond “long” more often than “short”. One way to determine if sub-arithmetic bisection represents a true bias favoring responding “long” is to extrapolate the best linear fit of bisection back to a log spread of 0 (TL = TS). In doing so we find that bisection is extremely close to the arithmetic mean ($B = (0.973 \pm 0.009) * AM$), indicating that subjects on average have almost no inherent bias to respond one way or the other.

Bisection as a function of absolute duration

While it is evident that the bisection point varies as a function of the *relative* distance between the reference durations, the spread, the questions remains as to whether the bisection point varies as a function of *absolute* duration. A number of studies on temporal bisection utilized a constant 4-fold spread while varying the length of the reference durations (L. G. Allan, 2002; Droit-Volet, 2003; Droit-Volet & Rattat, 2007; Droit-Volet et al., 2004; Droit-Volet & Wearden, 2001; Nichelli et al., 1996; Ortega & Lopez, 2008; Wearden, 1991; Wearden & Ferrara, 1995, 1996; Wearden et al., 1997). In this pooled data set the shortest TS is 100ms while the longest TL is 32s. In figure 2D we plot the bisection point, normalized by the arithmetic mean of the reference durations, as a function of the arithmetic mean of the reference durations. Analysis of this data shows that for the range of durations included here, bisection does not significantly vary as a function of absolute durations ($r = -0.156$, $p = 0.38$). For a given spread, bisection will occur at a constant fraction of the arithmetic mean of the reference durations (~94%), regardless of absolute duration.

Weber Ratio

Aside from the bisection point itself, it is also possible to determine the degree of discriminability the subject uses to parse the probe trials into the “short” and “long” categories, from the psychometric function. This metric, referred to as the Weber Ratio or Weber Fraction (R. H. Brown, 1960; Gibbon, 1977; Hirsh, Monahan, Grant, & Singh, 1990; Hobson, 1975), is defined according to the following equation: $(T(PL=0.75) - T(PL=0.25)) / T(PL=0.5)$; where $T(PL=X)$ is the duration at which a subject responded “long” with probability X. The duration $T(PL=0.5)$ is the bisection point, and the difference in durations $T(PL=0.75) - T(PL=0.25)$ is defined as the “just noticeable difference”—i.e., the smallest change in the stimulus that produces a substantial change in behavior (Figure 1B). A subject with a high degree of discriminability would produce a psychometric curve that appears very step-like, resulting in a low Weber Ratio, while another subject with a poorer discriminability would produce a more gradual psychometric function, resulting in a higher Weber Ratio (Figure 1B).

Weber ratio changes with spread

Individual studies have reported previously that the Weber Ratio decreases as the spread between the reference durations decreases (Nichelli et al., 1996; Siegel, 1986; Siegel & Church, 1984; Wearden, 1991; Wearden et al., 1997); a finding that was first observed with rats (Siegel & Church, 1984). That is, subjects discriminate better, leading to a steeper psychometric function, when the reference durations are pushed together. Here we show that this observation remains true across the much larger, multi-study, data set (L.G. Allan & Gibbon, 1991; Droit-Volet, 2003; Droit-Volet & Rattat, 2007; Droit-Volet et al., 2004; Droit-Volet & Wearden, 2001; Nichelli et al., 1996; Ortega & Lopez, 2008; Wearden, 1991; Wearden et al., 1997). The Weber Ratio does show a positive correlation with reference duration spread (Figure 3A, $r = 0.323$, $p = 0.04$). This result does not violate the scalar timing hypothesis which predicts a constant Weber Ratio across durations (Fetterman & Killeen, 1992; Gibbon, 1977). If we restrict our analysis to experiments using the same

spread, and focus on spread = 4 (Droit-Volet, 2003; Droit-Volet & Rattat, 2007; Droit-Volet et al., 2004; Droit-Volet & Wearden, 2001; Nichelli et al., 1996; Ortega & Lopez, 2008; Wearden, 1991; Wearden et al., 1997) since it yields the most data points and the largest range of absolute durations, we see that Weber Ratio shows no significant correlation with the arithmetic mean of the reference durations (Figure 3B, $r = 0.017$, $p = 0.87$). The question still remains why the Weber Ratio should vary as a function of reference duration spread. One possible explanation would be that in most experiments subjects are pre-trained with performance feedback on the reference durations prior to being presented with the probe durations. Learning reference durations close together (low spread) would then provide more feedback for stimuli close to the bisection point, and lead to a steeper psychometric function (low Weber Ratio) than learning with reference durations far apart, for which feedback is provided only far from the bisection point. Unfortunately, the data set does not contain Weber Ratios across a range of spreads with and without pre-training to allow us to test this hypothesis.

Effect of probe spacing

We have already seen that the reference durations used affect a subject's bisection point. The choice of probe durations also has a significant effect. Raslear (Raslear, 1983), in a study involving rats, first reported the observation that the bisection point is reduced if the probe durations are spaced logarithmically compared to the bisection point when the probe duration are spaced linearly between the same reference durations. Wearden and Ferrara (Wearden & Ferrara, 1995) later observed this phenomenon with human subjects. The multi-study combined data set contains 16 pairs of experiments where all conditions were constant within a pair except one experiment used logarithmically spaced probes and the other used linear spaced probes (L. G. Allan, 2002; L.G. Allan & Gibbon, 1991; Wearden & Ferrara, 1995, 1996; Wearden et al., 1997). 15 out of 16 experiment pairs show either reduced or equivalent bisection in the log-spaced probes experiment compared with the linear-spaced probes experiment ($p < 0.01$, Wilcoxon Rank Test). Further analysis reveals that the effect of logarithmic versus linear probe trials increases as the spread of the reference durations increases. In figure 4A we plot the ratio of the bisection points (logarithmic probe trials / linear probe trials) as a function of the spread of the reference durations. This clearly shows a negative correlation ($r = -0.906$, $p < 0.01$): at greater duration spreads, logarithmic probe spacing produces a greater reduction in the bisection point. In contrast to this effect of spread, absolute duration appears to have little impact. Figure 4B shows an analysis of the 5 experiment pairs that use a constant reference duration spread (spread = 4, shortest TS = 100ms, longest TL = 8s) (L. G. Allan, 2002; Wearden & Ferrara, 1995; Wearden et al., 1997) but investigate different absolute durations. No significant effect of absolute duration was found (Figure 4B, $r = -0.27$, $p = 0.8$).

Only one study to date (G. D. Brown, McCormack, Smith, & Stewart, 2005) has used probe spacings other than linear or logarithmic. In their 2005 study Brown et al. used 4 probe spacing regimes: superlogarithmic (durations are spaced uniformly in double log space), logarithmic, linear, and anti-logarithmic (logarithmic distribution reflected around the arithmetic mean). They found that bisection was lowest for the superlog probes, higher for the log probes, even higher for the linear probes, and highest for the anti-log probes. However, because no other studies used these probe spacing regimes, and because a range of absolute durations and spreads were not examined, we can perform no additional meta-analysis.

As we saw earlier, bisection with linearly spaced probes is subarithmetic. Performing the same analysis for logarithmically spaced probes (L. G. Allan, 2002; L.G. Allan & Gibbon, 1991; Lieving et al., 2006; Penney et al., 2000; Penney et al., 2005; Wearden & Ferrara, 1995, 1996; Wearden et al., 1997) we see that bisection is generally more subarithmetic

compared with linear spaced probe trials (Figure 2C), but still shows a significant trend to become more subarithmetic as reference duration spread increases (Figure 5A, $r = -0.841$, $p < 0.01$). Extrapolating the linear fit to the y-intercept ($B = 1.056 \pm 0.017 \cdot AM$) again shows subjects have almost no inherent bias to respond one way or the other. Finally, bisection with logarithmically spaced probe trials remains suprageometric at spreads greater than or equal to 2 (Figure 5B), similar to that for linear spaced probe trials (Figure 2B).

Areas of the data set requiring further study

Despite the enormous number of well crafted experiments and studies that went into creating this data set there are still areas where further experimentation would prove invaluable. Sylvie Droit-Volet and colleagues have contributed a number of insightful studies looking at the effect of age on temporal bisection (Droit-Volet, 2003; Droit-Volet et al., 2007; Droit-Volet & Rattat, 2007; Droit-Volet et al., 2004; Droit-Volet & Wearden, 2001). While the data explores a range of absolute durations (shortest TS = 200ms and longest TL = 8s), there are currently no experiments testing the effect of reference duration spread or probe duration spacing (all studies to date use a spread of 4 and linearly spaced probes) across children of different ages. Learning how the many peculiarities of human performance on this task develop with age will be essential to fully understanding how the neural mechanisms for time perception develop with age.

As discussed earlier, the Weber Ratio is a metric that quantifies a subject's degree of discriminability across the probe trials. Unfortunately only about half of the studies in the data set reported the Weber Ratio for their experiments. In particular it would be useful to see how the Weber Ratio changes with spread in experiments with and without pre-training on the reference durations. Finally, all data sets can benefit by expansion of parameters beyond the range currently explored. The range of absolute durations currently explored runs from 100ms for TS to 32s for TL while the range of reference duration spreads runs from 1.2 to 19. Our hope is that by compiling this multi-study data set together, its strengths and weaknesses will become apparent, thus helping to focus future studies and increasing our understanding of how time is processed in the brain.

Modeling the Temporal Bisection Task

To date a diverse array of models have been proposed to explain human performance on the temporal bisection task (L. G. Allan, 2002; L. G. Allan & Gerhardt, 2001; L.G. Allan & Gibbon, 1991; G. D. Brown et al., 2005; Droit-Volet et al., 2007; Droit-Volet et al., 2004; Droit-Volet & Wearden, 2001; Penney, Allan, Meck, & Gibbon, 1998; Penney et al., 2000; Wearden, 1991; Wearden & Ferrara, 1995, 1996). None however succeed in producing all the peculiarities discussed earlier. To briefly review, a successful model must produce all ten of the following phenomena. When probe trials are spaced linearly bisection will 1) be subarithmetic, 2) become more subarithmetic with increasing spread, and 3) extrapolate to the arithmetic mean when the spread equals 1. When the probe trials are spaced logarithmically bisection will be 4) more subarithmetic than with linear probes, 5) overall supra-geometric, and 6) more supra-geometric as spread increases. Irrespective of probe spacing bisection will 7) not be affected by absolute duration. When comparing bisection for logarithmically and linearly spaced probe trials bisection will be 8) lower with logarithmically spaced probes, and 9) the difference increases with increasing spreads. Finally, 10) the Weber Ratio decreases with decreasing spreads. While it is possible to design a model with a large number of free parameters to specifically address each phenomena, doing so would offer no potential insight into their root cause. Here we seek a model that will reproduce each of these phenomena with only a few (intuitively motivated) free parameters.

As with other models, we assume the subject has already learned the two reference durations. Previous work has shown that memories of durations are inherently noisy and can be modeled as scalar Gaussian distributions (Figure 6)(Church & Gibbon, 1982;Fetterman & Killeen, 1992;Gibbon, 1977), meaning they have a constant coefficient of variation. The probability that a presented duration “d” will be recognized as the learned duration “D” is given by the height of the Gaussian distribution at “d”. Similar ideas have been presented previously as a likelihood (Gibbon, 1981) or proximity rule (Siegel, 1986). Such a model for the memories of the reference durations leads to an obvious prediction: when the reference durations are spaced well apart (Figure 6a), there will be a wide range of intermediate durations that will be recognized as such by the subject, but as the reference durations are moved closer together (Figure 6b) subjects will find it more difficult to recognize intermediate durations as being a distinct group. Even though subjects are clearly instructed to indicate which reference duration they feel the presented duration is closest to, if they feel it *is* one of the two reference durations, no relative comparison should be required.

During the course of an experiment, a subject will receive tens to hundreds of trials with no feedback (as this is a categorization task, there are no right or wrong answers). We know humans are not ideal agents in these situations, and so we will draw on their biases to motivate some features of our model. One such bias that has been previously studied is commonly known as the “Gambler’s Fallacy”(Caruso, Waytz, & Epley, 2010; Jarvik, 1951; Nicks, 1959; Tversky & Kahneman, 1971). Here the subject assumes that the distribution of outcomes over the short-term should be equal to the distribution of trials over the long-term. Therefore, if the subject happens to make multiple similar responses in a row, e.g. responding “short” five times consecutively, they will be biased to expect and therefore favor responding oppositely, e.g. respond “long”, over the next few trials.

Two-step decision model

Here we present a simple two-step decision model based on the principles discussed above that contains only two free parameters and is capable of reproducing all the peculiarities in the human data identified earlier. Early work has demonstrated that perceived time is proportional to real time (L. G. Allan, 1979; Gibbon, 1981; Gibbon & Church, 1981). Here we assume perceived time is equal to real time. The model uses scalar Gaussian distributions to represent the noisy memories of the two reference durations as discussed above (Figure 6). In the first step the subject determines whether the stimulus is (a) the short reference duration, at which point they would respond “short”, (b) the long reference duration, at which point they would respond “long”, or (c) an intermediate duration. If the stimulus is determined to be an intermediate duration then the decision process moves to step 2. Here the subject compares the relative distance between the stimulus and both reference durations, and responds according to whichever is closer (Figure 6a). Previous responses bias future choices in accordance with the Gambler’s Fallacy. Each of the two steps is explained in more detail below.

Step 1. Here we model the memory of each of the two reference durations as a Gaussian distribution over durations, with a mean equal to the reference duration, and a standard deviation proportional to the reference duration (Figure 6). We then take the probability of a stimulus duration being labeled as one of the reference durations as equal to the reference memory density distribution, evaluated at the stimulus duration. The distributions range from 0 to 1, meaning a subject is 100% likely to identify each reference duration correctly. There are four possible outcomes to this decision step: 1) The stimulus is determined to be TS, respond “short”; 2) The stimulus is determined to be TL, respond “long”; 3) The stimulus is determined to be both TS and TL, a guess is made with equal probability of responding “short” or “long”; 4) The stimulus is determined to be neither TS nor TL, then it must be an intermediate probe duration, continue to step 2. This step is similar to the

“likelihood” (Gibbon, 1981) and “proximity rule” (Siegel, 1986; Siegel & Church, 1984) models proposed earlier.

Step 2. This step is similar to the decision model first presented by Gibbon and Church (Gibbon, 1981; Gibbon & Church, 1981) and later modified (Penney et al., 1998; Raslear, 1983; Wearden, 1991). Here the subject compares the stimulus duration “s” to each reference duration stored in memory. Whichever duration the stimulus is closer to is the subject’s response. The reference duration values pulled from memory are drawn from the scalar Gaussian distributions used to model them (Figure 6b). One value t_s is drawn from the TS distribution, and one value t_l is drawn from the TL distribution. During this step the subject is biased by their previous decisions, meaning if they make many “long” responses they will be biased to respond “short” and vice versa. Such a bias is based on the “gambler’s fallacy” described previously. The model works from the prior expectation that “long” and “short” responses should be produced with equal likelihood. The bias factor “b” starts at 1 (no bias) and evolves during the session. If $\text{abs}(t_s - s) * b < \text{abs}(t_l - s)$ then the subject responds “short”, if not, the subject responds “long”. After each “short” response b is increased by some factor “x” [$b = b * x$, where x must be greater than 1] while after each “long” response b is reduced [$b = b / x$]. Equilibrium is reached when the subject produces an equal number of “long” and “short” responses, i.e. their expectation is satisfied.

Free parameter fitting and model performance

This model contains only two free parameters, the CV of the TS and TL distributions stored in memory, and x, the rate at which the bias changes. The parameters were fit by testing the full parameter space over a range of values (CV range: 0 to 1, resolution 0.01; X range: 1 to 1.3, resolution 0.005). The output of the model is an average of 100 runs each consisting of 200 trials for each spread value represented in the human data set. At the start of each run the bias factor, b, is reset to 1. For each parameter combination five fits were calculated (Linear Probes Bisection v. Spread, Log Probes Bisection v. Spread, Linear Probes Weber v. Spread, Log Probes Weber v. Spread, and Linear / Log Probe Bisection v. Spread). A score was assigned to each fit according to Equation 1.

$$\text{Score} = 1 - (\text{sum}((\text{model} - \text{data})^2) / \text{sum}(\text{data}^2)) \quad \text{Equation 1}$$

The overall score for a particular parameter combination was the product of the five individual scores. The fitness landscape contains a broad region of parameter values that give equivalently good fits. The model output with parameter values CV = 0.16 and x = 1.05 produced the best fit and is shown in Figure 7. Here the model shows a statistically significant fit to the human performance data (individual correlations between model and human data shown in Figure 7 subpanels). The correlation between the model and bisection normalized by the arithmetic mean (Figure 7A), and the model and the Weber Ratio (Figure 7G) are both equivalent to the correlation between that same data and the best linear fits to that data (Bisection/AM: $r = 0.2$, $p = 0.061$; Weber: $r = 0.32$, $p = 0.045$).

How the model accounts for each facet of performance

The two-step nature of the decision model is key in its ability to model human performance on the temporal bisection task. Since the reference duration memories are scalar, the TL distribution will always be wider than the TS distribution. This establishes a slight asymmetry in the durations that are determined to be neither TS nor TL and therefore proceed to step 2. For a uniformly sampled range of durations step 2 in the decision model will naturally produce arithmetic bisection. However, because of the asymmetry established in step 1, where the durations are shifted towards the short reference duration, more “short”

responses will be elicited than “long” responses. Our implementation of the Gambler’s Fallacy assumes a uniform prior distribution, i.e. there should be an equal number of “short” and “long” responses. Therefore, the bias term will compensate by shifting bisection towards shorter durations, i.e. the model will bisect subarithmetically. As we saw earlier, such subarithmetic bisection will produce a transition from arithmetic to geometric to sub geometric bisection as the spread between the reference durations is reduced. One alternative which has been proposed previously is to simply add a “choose short” bias (Wearden, 1991). While this will produce subarithmetic bisection, such a bias will be constant over reference duration spreads, and will still be apparent when TS equals TL. This model produces bisection that shows no bias when TS equals TL and becomes more subarithmetic as the reference durations spread increases (Figure 7A), in agreement with the human data.

As the reference durations are moved closer together, the range of durations that require step 2 in the decision model is decreased. This is because when the Gaussian functions representing TS and TL in memory overlap more and more, the chance of a stimulus not being categorized as TS or TL in step 1 goes down. A model based solely on step 1 will produce bisection with a steeper psychometric function (have a lower Weber Ratio) than one built solely on step 2. Therefore, as the spread between the reference durations is decreased, the range of durations requiring step 2 is decreased, and the overall psychometric function becomes steeper (the Weber Ratio goes down).

Finally, the model will produce a lower bisection value for logarithmically spaced probe trials than for linearly spaced probe trials because of the bias change parameter x used in step 2. In the case of log spaced probe trials, the model initially will bisect arithmetically. However, since more trials are presented below the arithmetic mean than above, the model will respond “short” more than “long”. This causes the bias factor b to be increased more than decreased, thus causing the bisection point to drop. As the reference durations are moved apart, the distance between the geometric and arithmetic means increases. Therefore, larger spreads lead to greater differences in bisection for linear and log probe trials. When applied to more extreme spacings of probe trials such as the anti-logarithmic and super-logarithmic spacings found in (G. D. Brown et al., 2005), the model produces bisection values equivalent to those of the human subjects in that study.

Model predictions

Nothing about this 2-step decision model requires it to be exclusively applicable to decisions involving durations. The bisection task can be performed relative to other stimulus dimensions such as frequency. If this model is in fact a general decision model, then it predicts that the same peculiarities as seen in the temporal bisection task will appear in other bisection tasks. Recently Gordon Brown and colleagues (G. D. Brown et al., 2005) performed a frequency bisection task for human subjects. Interestingly, subjects showed subarithmetic bisection that was closer to the geometric mean when the reference frequency spread was small. Also, switching from linear to log spaced probe trials had a greater effect on the bisection point for the larger reference frequency spread, and the Weber Ratio was reduced when the reference frequency spread was reduced. All of these results are predicted by our model when used as a general decision model for bisection tasks.

For logarithmically spaced probe trials the model predicts that subjects should start off bisecting arithmetically ($b = 1$) and within a few trials quickly shift towards more geometric-like bisection (as b is increased repeatedly by x). With trial-by-trial performance of enough subjects performing the task, such a shift should be detectable.

The canonical version of the temporal bisection task only includes probe durations that are intermediate to the reference durations. One study of rats performing this task (Siegel & Church, 1984) has shown that responses reverse outside the reference duration range. That is, the fraction of trials resulting in a “long” report *increases* as probe durations become shorter than TS, and *decreases* as probe durations become longer than TL. In contrast, our model produces a monotonic psychometric function both within and beyond the reference durations. The model therefore predicts that this finding from rats will not be replicated in humans.

Finally, the implementation of the Gambler’s Fallacy in step 2 works off a uniform expectation of responding “short” and “long”. One means to test if such an expectation does exist and can influence performance would be to instruct participants prior to beginning the experiment that the distribution of intermediate durations will not be uniform. In one experiment participants could be told there will be more durations near the short reference duration, while in another they can be told the opposite. However, due to the pervasive fundamental nature of this fallacy, subjects might not be able to ignore it, and therefore another test would be to simply analyze trial-by-trial performance, looking for an affect of previous responses on future choices.

Possible alternative models

The key feature of this model is that the decision is made in a two-step process: the first step determining if the stimulus is either of the reference durations, and the second step comparing the relative distance between the stimulus and reference durations on a trial to trial bias. Precisely how these steps are implemented mathematically is less important for obtaining a good fit to the behavioral data. We tested models that used either an absolute threshold in step 1 (i.e. if the stimulus was within say 1 standard deviation of the reference duration it was determined to be the reference duration) or a combination of absolute thresholds and probabilistic decisions and obtained reasonably equivalent fits. We also tested incrementing the bias in step 2 in a linear manner rather than the multiplicative manner described above and obtained equivalent fits. While we utilized the Gambler’s Fallacy to implement the bias in step 2, any mechanism that produces an equivalent bias will fit the data equally well. We also tested a number of alternative behaviors the model can use during step 1 when the stimulus is determined to be both the short and the long reference durations. Making an unbiased guess (implemented above), iterating step 1 until the stimulus is determined to not be both reference durations, or moving onto step 2 all produce equivalent fits. However, in reducing the model to a simple 1-step decision mechanism we were unable to reproduce all of the peculiarities present in the human data. Also, updating and applying the bias term to both steps 1 and 2, only step 1, or omitting it entirely produced significantly worse fits.

Model shortcomings

This model is purely a decision model and does not take into account many other steps that are required to perform the temporal bisection task. These include learning the reference durations, any biases that may be imposed upon those memories as trials progress, and the recall of reference duration values from long-term memory and transfer to short-term memory. Some studies have shown that durations held in memory are subject to decay over a time course of seconds as evidenced by the “choose short” effect (Lieving et al., 2006). Others have shown that presenting the reference durations before an intermediate probe duration can induce different biases depending on the order of reference duration presentation (TS then TL then probe vs. TL then TS then probe) (L. G. Allan & Gerhardt, 2001). Still others have shown that auditory and visual stimuli may be measured with neural timers that run at different rates (Penney et al., 1998). When both modalities are interleaved

in the same session their memories can mix, causing a performance bias. An early work by Warren Meck (Meck, 1983) showed that many of these steps can be pharmacologically dissociated, and therefore can likely be modeled separately with the output of one model, e.g. measuring durations, being the input to another model, e.g. storing durations in memory.

DISCUSSION

Review of human idiosyncrasies

Here we have combined the data from 18 studies conducted over the past 19 years, each performing the temporal bisection task on human subjects, to amass a combined data set of roughly 1,020 individuals performing a total of over 302,000 trials. From this we were able to not only tease out many idiosyncrasies in human performance on this task, some of which have been reported previously, but also determine whether they are statistically significant.

First, while it is true that subjects do show an arithmetic to geometric to subgeometric transition in their bisection point as the reference durations are moved closer together, this is not because of some fundamental change in how time is measured or decisions are made. Instead, we have shown that what the data in fact shows is subarithmetic bisection across all reference duration spreads; with bisection becoming more subarithmetic as the reference durations are moved apart. Second, we have confirmed that a subject's Weber Ratio does indeed increase as the reference durations are moved apart. This is true for both linear and logarithmically spaced probe trials. Finally, we confirmed that not only do logarithmically spaced probe trials decrease the bisection point relative to linearly spaced probe trials, but the absolute amount the bisection point is shifted is positively correlated with the spread of the reference durations.

Antecedents to the Model

The 2-step model we have proposed here has its roots in two broad classes of models. Our step 1 is similar to likelihood rules (Gibbon, 1981; Siegel, 1986; Siegel & Church, 1984). Our step 2 is closely related to similarity rules (L.G. Allan & Gibbon, 1991; Gibbon, 1981; Penney et al., 1998; Pfanzagl, 1968; Wearden, 1991). Both of these types of rules have received much attention over at least the past three decades. Our modeling contribution is to combine these two rules into a single 2-step decision model. In fitting with the "likelihood" rule, during our step 1, subjects decide if the stimulus is either of the two reference durations. This rule typically fails when the reference durations are well spaced and subjects should be aware that the stimulus is clearly of an intermediate duration. For that reason we included a second step, based on the "similarity" rule, where subjects then compare the temporal distance between the stimulus and either of the reference durations. We found that by combining both rules into a single model we were able to explain a remarkably large number of features in the data, using only two free parameters. The model we propose also incorporates the concepts of scalar timing (Fetterman & Killeen, 1992; Gibbon, 1977), linear subjective time (L. G. Allan, 1979; Gibbon, 1981; Gibbon & Church, 1981), and prior expectancy (Caruso et al., 2010; Jarvik, 1951; Nicks, 1959; Tversky & Kahneman, 1971), all of which have been extensively studied in the literature.

Review of model performance

While each of the idiosyncrasies uncovered above may seem insignificant, in light of the large data set from which they were measures, each is both robust and statistically significant. Therefore any model designed to explain human performance on the temporal bisection task must account for each one. Here we have presented a novel 2-step decision model that not only accounts for all the peculiarities of human performance discussed above, but also makes predictions that will need to be tested in future experiments. Briefly, subjects

first decide if the stimulus is either of the two reference durations. Only if they conclude it is neither do they then determine which reference duration the stimulus is closest to. Not only is this simple model with only two free parameters able to fit the combined data set exceptionally well, it also makes intuitive sense. With reference durations that are well spaced apart, subjects are likely to recognize the intermediate probe durations as just what they are, and perform the task as described (report which reference duration the stimulus is more similar to, i.e. Step 2). However, when the reference durations are placed very near one another, subjects will likely not appreciate that the probe durations are distinct from the reference durations, and will therefore be more likely to report which reference duration they thought they heard or saw, i.e. Step 1.

The model we have presented makes no claims on the neural mechanism used to perform each of the steps. The intuition that a duration which involves both steps in the model should involve a decision that is slower and more difficult than a duration that involves only the first step need not be the case. One can imagine an implementation of our model with a drift diffusion mechanism. The more balanced the forces, e.g. when one is close to the psychometric middle, or when the spread of the reference durations is small, the longer the DDM will take to reach a bound thereby make a decision. Reaction times however are notoriously difficult to measure in a task involving temporal stimuli. Another possibility is that both step's decision processes are computed simultaneously rather than serially, with the decision from step 2 only being recognized if step 1 determines the stimulus to have an intermediate duration. Only through studies involving electrophysiological and pharmacological techniques will we begin to gain the knowledge required to produce more detailed mechanistic models.

The reference durations in our model receive special status, being required for both steps. We did this precisely because the vast majority of studies in the data set present the subjects with specific verbal instructions singling out the reference durations as the durations the probe stimuli should be compared with. Also, in most experiments the only feedback ever given is on the reference durations. Previous studies have demonstrated that subjects can categorize durations without explicit identification of the reference durations (Droit-Volet & Rattat, 2007; Wearden & Ferrara, 1995). While it may be possible to construct a decision model without special identification of the reference durations, such as removing step 1 and having the decision in step 2 be based around a central threshold (Wearden & Ferrara, 1995), our attempts to do so have not been successful in fitting all of the aspects of the data presented here. Given an expanded data set on this modified version of the temporal bisection task, we may find that performance is slightly altered compared with the canonical version presented here.

Variations on the model are possible

Precisely how the decision in step 1 is reached (is the stimulus TS, is it TL, is it neither, or is it both) is far less important than having a 2-step decision process in general. Here we reported that the probability of the model determining that a stimulus is one of the reference durations is given by the value of the Gaussian function modeling the memory of that reference duration at the stimulus duration. However, a number of variations on this rule produced similar overall results. These include using a hard fixed boundary to decide if a stimulus is a reference duration, such as ± 1 standard deviation of the reference duration memory, or using a probabilistic decision only within a hard fixed boundary. Also, in step 2, exactly how a subject's bias changes over time, is it a fractional change (as we modeled here), an absolute change, and does it only depend on the response on the previous trial or is it influenced more by the recent history of trials, remains to be determined. Importantly, all such variations involve a 2-step decision and produce very similar results that can only be teased apart with further experimentation.

What we hope will come of this

While the data set compiled here is truly immense in size, there are still a number of dimensions for which further experimentation would be fruitful. These include but are not limited to, exploring various probe spacings and reference duration spreads across children of different ages, testing reference duration spreads and absolute durations beyond the extremes of the data set, assaying how bisection varies across trials when probes are spaced non-linearly, measuring performance on durations that lie outside the reference duration range, and exploring how individuals with various psychiatric diseases differ from normal human performance. We hope that pooling the data together will focus future experiments to explore the more sparse regions of the data set.

The temporal bisection task involves measuring, learning, retrieval, storage, and comparison of durations, yet the task itself as routinely performed is only designed to analyze how durations are measured and compared. While some studies have been performed to explore how durations are stored in both long-term and short-term memory, the answer is still far from conclusive, and therefore ripe for further examination. A large, more comprehensive data set can only aid in producing more accurate models of how time is processed in our brains.

The data presented here has been exclusively from studies involving humans, yet the temporal bisection task has also been performed in a range of animal species. While interpreting their data will be extremely useful in developing a broader model of how time is processed in the brain, care must also be taken. Interpretation of responses in this task involves the assumption that the subjects are following the specific verbal instructions. As of yet no one has demonstrated the ability to deliver such instructions to non-human subjects. Therefore care must be taken when interpreting such results as one cannot assume what problem the animals are actually trying to solve.

Acknowledgments

We would explicitly like to thank all the authors of the studies that went into this data set, especially John Wearden, Lorraine Allan, Warren Meck, John Gibbon, Trevor Penney, and Sylvie Droit-Volet for their significant contributions of data and insight. We would also like to thank the members of the Brody Lab, especially Jeffrey Erlich, Bingni Bruton, Max Bialek, and Joseph Jun for helpful comments in preparing this manuscript.

REFERENCES

- Allan LG. The perception of time. *Perception & Psychophysics* 1979;26(5):340–354.
- Allan LG. The location and interpretation of the bisection point. *Q J Exp Psychol B* 2002;55(1):43–60. [PubMed: 11900306]
- Allan LG, Gerhardt K. Temporal bisection with trial referents. *Percept Psychophys* 2001;63(3):524–540. [PubMed: 11414139]
- Allan LG, Gibbon J. Human Bisection at the Geometric Mean. *Learning and Motivation* 1991;22:39–58.
- Bovet P. Subjective scales of duration obtained by a method of bisection. *Annee Psychol* 1968;68(1): 23–36. [PubMed: 5701175]
- Brown GD, McCormack T, Smith M, Stewart N. Identification and bisection of temporal durations and tone frequencies: common models for temporal and nontemporal stimuli. *J Exp Psychol Hum Percept Perform* 2005;31(5):919–938.
- Brown RH. Weber Ratio for Visual Discrimination of Velocity. *Science* 1960;131(3416):1809–1810. [PubMed: 17753209]
- Caruso EM, Waytz A, Epley N. The intentional mind and the hot hand: Perceiving intentions makes streaks seem likely to continue. *Cognition* 2010;116(1):149–153. [PubMed: 20472229]

- Church RM, Deluty MZ. Bisection of temporal intervals. *J Exp Psychol Anim Behav Process* 1977;3(3):216–228. [PubMed: 881613]
- Church RM, Gibbon J. Temporal generalization. *J Exp Psychol Anim Behav Process* 1982;8(2):165–186. [PubMed: 7069377]
- Droit-Volet S. Alerting attention and time perception in children. *J Exp Child Psychol* 2003;85(4):372–384. [PubMed: 12906848]
- Droit-Volet S, Meck WH, Penney TB. Sensory modality and time perception in children and adults. *Behav Processes* 2007;74(2):244–250. [PubMed: 17084041]
- Droit-Volet S, Rattat AC. A further analysis of time bisection behavior in children with and without reference memory: the similarity and the partition task. *Acta Psychol (Amst)* 2007;125(2):240–256. [PubMed: 17055990]
- Droit-Volet S, Tournet S, Wearden J. Perception of the duration of auditory and visual stimuli in children and adults. *Q J Exp Psychol A* 2004;57(5):797–818. [PubMed: 15204118]
- Droit-Volet S, Wearden JH. Temporal bisection in children. *J Exp Child Psychol* 2001;80(2):142–159. [PubMed: 11529672]
- Fetterman JG, Killeen PR. Time discrimination in *Columba livia* and *Homo sapiens*. *J Exp Psychol Anim Behav Process* 1992;18(1):80–94. [PubMed: 1578202]
- Gibbon J. Scalar Expectancy theory and Weber's law in animal timing. *Psychological Review* 1977;84:279–325.
- Gibbon J. On the form and location of the psychometric bisection function for time. *Journal of Mathematical Psychology* 1981;24:58–87.
- Gibbon J, Church RM. Time left: linear versus logarithmic subjective time. *J Exp Psychol Anim Behav Process* 1981;7(2):87–107. [PubMed: 7241054]
- Hirsh IJ, Monahan CB, Grant KW, Singh PG. Studies in auditory timing: 1. Simple patterns. *Percept Psychophys* 1990;47(3):215–226. [PubMed: 2326145]
- Hobson SL. Discriminability of fixed-ratio schedules for pigeons: effects of absolute ratio size. *J Exp Anal Behav* 1975;23(1):25–35. [PubMed: 16811828]
- Jarvik ME. Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *J Exp Psychol* 1951;41(4):291–297. [PubMed: 14850645]
- Killeen, PR.; Fetterman, JG.; Bizo, LA. Time's Causes. In: Bradshaw, CM.; Szabadi, E., editors. *Time and Behavior: Psychological and Neurological Analyses*. Amsterdam: Elsevier; 1997. p. 79-131.
- Lieving LM, Lane SD, Cherek DR, Tcheremissine OV. Effects of delays on human performance on a temporal discrimination procedure: evidence of a choose-short effect. *Behav Processes* 2006;71(2–3):135–143. [PubMed: 16289517]
- Meck WH. Selective adjustment of the speed of internal clock and memory processes. *J Exp Psychol Anim Behav Process* 1983;9(2):171–201. [PubMed: 6842136]
- Nichelli P, Alway D, Grafman J. Perceptual timing in cerebellar degeneration. *Neuropsychologia* 1996;34(9):863–871. [PubMed: 8822733]
- Nicks DC. Prediction of sequential two-choice decisions from event runs. *J Exp Psychol* 1959;57(2):105–114. [PubMed: 13641581]
- Ortega L, Lopez F. Effects of visual flicker on subjective time in a temporal bisection task. *Behav Processes* 2008;78(3):380–386. [PubMed: 18358636]
- Penney, TB.; Allan, LG.; Meck, WH.; Gibbon, J. Memory Mixing in Duration Bisection. In: Rosenbaum, DA.; Collyer, CE., editors. *Timing of Behavior: Neural, Psychological and Computations Perspectives*. Cambridge, MA: MIT Press; 1998. p. 165-193.
- Penney TB, Gibbon J, Meck WH. Differential effects of auditory and visual signals on clock speed and temporal memory. *J Exp Psychol Hum Percept Perform* 2000;26(6):1770–1787. [PubMed: 11129373]
- Penney TB, Meck WH, Roberts SA, Gibbon J, Erlenmeyer-Kimling L. Interval-timing deficits in individuals at high risk for schizophrenia. *Brain Cogn* 2005;58(1):109–118. [PubMed: 15878731]
- Pfanzagl, J. *Theory of measurement (in cooperation with V. Baumann and H. Huber)*. New York: Wiley; 1968.

- Raslear TG. A test of the Pfanzagl bisection model in rats. *J Exp Psychol Anim Behav Process* 1983;9(1):49–62. [PubMed: 6842132]
- Siegel SF. A test of the similarity rule model of temporal bisection. *Learning and Motivation* 1986;17:59–75.
- Siegel SF, Church RM. The decision rule in temporal bisection. *Annals New York Academy of Sciences* 1984;423:643–645.
- Tversky A, Kahneman D. Belief in the law of small numbers. *Psychological Bulletin* 1971;76:105–110.
- Wearden JH. Human performance on an analogue of an interval bisection task. *Q J Exp Psychol B* 1991;43(1):59–81. [PubMed: 2017575]
- Wearden JH, Ferrara A. Stimulus spacing effects in temporal bisection by humans. *Q J Exp Psychol B* 1995;48(4):289–310. [PubMed: 8532899]
- Wearden JH, Ferrara A. Stimulus range effects in temporal bisection by humans. *Q J Exp Psychol B* 1996;49(1):24–44. [PubMed: 8901385]
- Wearden JH, Rogers P, Thomas R. Temporal bisection in humans with longer stimulus durations. *Q J Exp Psychol B* 1997;50(1):79–94. [PubMed: 9141912]

Figure 1.

The Temporal Bisection Task. A) Training is split into 2 phases. During the Pre-Training Phase subjects learn the reference durations and receive feedback. During the Probe Duration Phase intermediate durations are added and feedback is withheld. Subjects are instructed to indicate which reference duration they feel the stimulus duration is closest too. B) Probability of a subject responding "long" as a function of stimulus duration. The bisection point is defined as the duration which produces 50% "long" responses.

Figure 2.

Bisection with linearly spaced probe trials. A) Log ratio of the distance between the bisection point, B, and the geometric mean, GM, and the bisection point and the arithmetic mean, AM, as a function of spread (ratio of the reference durations). Values above zero are closer to the arithmetic mean, values below zero are closer to the geometric mean. B) Bisection normalized by the geometric mean of the reference durations as a function of spread. Regions marked by the gray box enlarged in the insert. Note that bisection is sub-geometric below $\text{log}_2(\text{spread})$ of 1. C) Bisection normalized by the arithmetic mean of the reference durations as a function of spread. Note that bisection becomes more sub-arithmetic with increasing spread. D) Bisection normalized by the arithmetic mean of the reference durations as a function of the arithmetic mean of the reference durations. Only experiments with a spread of 4 were used here. Bisection shows no correlation with absolute durations.

Figure 3.

Weber Ratio is affected by reference duration spread. A) Weber Ratio as a function of reference duration spread. Subjects show an increased discriminability (lower Weber Ratio) with decreasing spread. B) Weber Ratio as a function of the arithmetic mean of the reference durations. Only experiments with a spread of 4 were used here. Weber ratio shows no correlation with absolute duration. All experiments here used linearly spaced probe durations.

Figure 4.

Logarithmically vs. Linearly spaced probe durations. Each data point represents a pair of experiments where all conditions were constant except in one logarithmically spaced probe duration and in the other linearly spaced probe durations were used. A) Ratio of bisection points as a function of spread. The ability of logarithmically spaced probe durations to decrease the bisection point increases with increasing spread. B) Ratio of bisection points as a function of the arithmetic mean of the reference durations. Only experiments with a spread of 4 were used here. The effect of probe spacing is not influenced by absolute duration.

Figure 5.

Bisection with logarithmically spaced probe durations. A) Bisection point normalized by the arithmetic mean of the reference durations as a function of spread. Bisection becomes more subarithmetic as spread is increased. B) Bisection point normalized by the geometric mean of the reference durations as a function of spread. Bisection becomes more supra-geometric as spread is increased.

Figure 6.

2-Step Decision Model. A) Step 1: Probability of a subject deciding that a stimulus is either of the reference durations TS or TL is given by the TS and TL distributions respectively. Green curve shows the probability of a duration being recognized as neither reference duration and therefore moving on to step 2. B) Step 2: One value, t_s , is pulled from the TS distribution, and t_l is pulled from the TL distribution. The distance between these values and the stimulus, s , are compared. If $\Delta t_s < \Delta t_l$ the subject responds “short”, else the subject responds “long”. Note, as the spread is decreased (spread = 8 in A, spread = 4 in B) the probability of durations moving to step 2 is decreased.

Figure 7.

2-step decision model fit to human performance on the Temporal Bisection Task. Model fit shown in green, error bars are standard deviation across 500 independent runs each consisting of 200 trials. Black dots represent individual human experiments. Red dots are the mean of the human data for a particular spread, error bars standard deviation. A–C) Bisection as a function of spread with linearly spaced probe durations. D–F) Bisection as a function of spread with logarithmically spaced probe durations. G) Weber Ratio as a function of spread with linearly spaced probe durations. H) Weber Ratio as a function of spread with logarithmically spaced probe durations. I) Ratio of bisection with logarithmically and linearly spaced probe durations.