Running head: LOG VERSUS LINEAR TIMING

Log versus linear timing in human temporal bisection: Bringing the right answer to the wrong question

Jérémie Jozefowiez

Université de Lille & Universidade do Minho

Clément Gaudichon & Francis Mekkass

Université de Lille

Armando Machado

Universidade do Minho

Correspond with:
Dr. Jérémie Jozefowiez
Laboratory of Affective and Cognitive Sciences (SCALab)
Université de Lille
Campus de Lille 3
Domaine Universitaire du Pont de Bois
B.P. 60149
59653 Villeneuve d'Ascq Cedex
France
Email: jeremie.jozefowiez@univ-lille3.fr
Phone: +33-320-041-6866
Fax: +33-320-041-6036

Submitted: October 2017

Abstract

Keywords:

Log versus linear timing in human temporal bisection: Bringing the right answer to the wrong question

Interval timing is the ability of organisms to perceive duration in the second to minute range between two events. It is thought to be critical to adaptive behavior as it allows animals to learn the temporal structure of their environment (when does something happen relative to something else?).

Many computational models have been proposed to account for interval timing. The dominant model is the scalar expectancy theory (SET. Gibbon, Church, & Meck, 1984). The model assumes that timing the interval between two stimulus events relies on an internal pacemaker. The pacemaker starts emitting pulses at the onset of the first stimulus event (called the time marker), which will accumulate in short-term memory (STM). When the second event occurs, the number of pulses accumulated is stored in long-term memory (LTM) to serve as reference. To decide whether it is time for the second event to occur, the model compares its STM representation of the time elapsed since the onset of the time marker and its LTM representation of the time at which the second event occurs relative to the time-marker.

A fundamental assumption of SET is that time is represented linearly: Both the STM and the LTM representation of an interval t units of time long is a linear function of t. More precisely, SET assumes that the representation x of an interval t is random variable drawn from a Gaussian distribution with mean proportional to t and standard deviation proportional to the mean. This assumption regarding the standard deviation is necessary in order for the model to account for Weber's law, which is ubiquitous in interval timing (Lejeune & Wearden, 2006; Wearden & Lejeune, 2008).

An alternative compatible with Weber's law (for instance, see Falmagne, 2002) would be to assume that time is encoded logarithmically, that is to say that the representation of x of an interval t is random variable drawn from a Gaussian distribution with mean proportional to *ln t* and constant standard deviation. Such a representation scheme is not consistent with a pacemaker-driven mechanism, which entails a linear representation. It is, for instance, used by the Behavioral Economic Model (BEM. Jozefowiez, Staddon, & Cerutti, 2009). BEM is an associative model assuming that timing relies on associations which build between time-dependent states and responding.

Because of the critical importance of the linear timing assumption to SET, many attempts have been made to test whether time is indeed represented linearly but no definitive conclusion has been reached. For instance, Church and Deluty (1977) concluded in favor of log timing but Gibbon (1981) showed that their data could be accounted by SET. The same way, while Gibbon and Church (1981) concluded in favor of linear timing, several authors have shown that their results could also be explained by a model assuming log timing (for instance, Jozefowiez & Machado, 2013).

Yi (2009) proposed a new approach to that problem based on an analysis of bisection procedures with signal detection theory (SDT. See Wickens, 2002 for an introduction to SDT). A bisection procedure (Church & Deluty, 1977; Stubbs, 1968) is a temporal discrimination task where the subject is taught to emit response *r1* following a stimulus lasting *d1* units of time and response *r2* following a stimulus lasting *d2* units of time (*d1*<*d2*). A SDT analysis of this type of task assumes that the subject emits *r2* when *x*, its perception of the stimulus duration (what could be called the subjective stimulus duration), falls above a critical value (threshold) λ. When a stimulus lasting *d1* units of time is shown, *x* is drawn from a Gaussian distribution with mean μ_1 and standard deviation σ_1. When a stimulus lasting *d2* units of time is shown, *x* is drawn from a Gaussian distribution with mean μ_2 and standard deviation σ_2. Various techniques allow to retrieve allow to retrieve the parameters μ_1, μ_2, σ_1, and σ_2 of the SDT model from the data (see Wickens, 2002). Applying those data to data obtained from rats performing a bisection procedure, Yi (2009) concluded that the standard deviation of subjective time did not increase with its mean but rather seemed to remain constant, a result more consistent with log timing than with SET.

The implications of Yi's (2009) results for SET are not as straightforward as one might think. She did not demonstrate that rats encode time logarithmically. What she showed is that, if a model was mathematically equivalent to a Gaussian SDT model, it would be a better fit to her data if it assumes a constant standard deviation of the subjective stimulus duration than if it assumes that it is a linear function of the duration represented. An associative model like BEM is mathematically equivalent to a Gaussian SDT model (Jozefowiez & Machado, 2013). As it uses a log representation of time with constant variance, it is consistent with Yi's (2009) data. The case of SET is a bit more complicated, mainly because SET is not really a model, but more a general framework within which several models can be designed, which is why it has been so difficult to refute. Notably, different versions of SET can be built depending on whether noise is assumed in STM, LTM or in both and depending on the response rule used to generate behavior based on STM and LTM representations of stimulus durations.

When applied to a bisection task, all versions of SET assume that the participant compares a STM representation *f(t)* of the stimulus duration to LTM representations *f(S)* and *f(L)* of the duration of the short-duration stimulus and of the long-duration stimulus, respectively. The original version of the model (Gibbon, 1981) assumes a ratio response rule. The stimulus duration is categorized as long if

$$\frac{f(t)}{f(S)} > \theta \frac{f(L)}{f(t)} \qquad (1)$$

where θ is a bias parameter. This version of SET predicts a bisection point at the geometric mean. The full model assumes noise in both STM and LTM representations, but is rarely used because it is difficult to track computationally (Jozefowiez & Machado, 2013. See also Appendix I). It cannot be reduced to a Gaussian SDT model and simulations by Jozefowiez & Machado (2013) shows that it is actually mimics a SDT model with constant standard deviation. If noise is assumed only in STM, as it is often the case (for instance, see Allan, 2002a), the model can be reduced to a Gaussian SDT model and cannot account for Yi's (2009) data as it assumes that the noise in the representation of a duration increases with the value of that duration (see Jozefowiez & Machado, 2013).

Other versions of SET can be built if we change assumption regarding the response rule. Of particular note is a version of SET proposed by Wearden (1991) and which a difference decision rule. The stimulus duration is categorized as long if

$$f(t) - f(S) > \theta[f(L) - f(t)] \qquad (2)$$

As Appendix II shows, if noise is assumed only in STM, this version of SET can be reduced to a Gaussian SDT model. Hence, this version is not compatible with Yi's (2009) data neither but, for reasons explained below, it is rarely used to fit non-human data. Surprisingly, as shown in Appendix II, the version assuming noise in both STM and LTM is fatally flawed and cannot account for performance in a bisection procedure.

Usually, in a bisection task, once the *d1* vs. *d2* discrimination is mastered, the subject is presented with durations ranging between *d1* and *d2* to see if they are subjectively closer to *d1* (in this case, the subject would emit *r1*) or to *d2* (in this case, the subject will emit *r2*). Of special interest is the so-called bisection point, which is the duration for which the subject is indifferent between *r1* and *r2.* This duration is supposed to be psychologically equidistant from *d1* and *d2*, which is why the subject cannot make up his mind. An intriguing difference between human and non-human animals is the location of the bisection point: It is located at the geometric mean between *d1* and *d2* in non-human animals (i.e. Church & Deluty, 1977; Stubbs, 1968), such as rats or pigeons, and at the arithmetic mean between *d1* and *d2* in humans (Allan, 2002b; Allan & Gerhardt, 2001; Wearden, 1991; Wearden & Ferrara, 1995, 1996. A few study reports a bisection point at the geometric mean, i.e. Allan & Gibbon, 1991; Provasi, Rattat, & Droit-Volet, 2011).

A bisection point at the geometric mean of the two anchor durations *d1* and *d2* is a fundamental prediction of a model like BEM, which assumes log timing. It is very difficult to see how such models could be modified in order to account for a bisection point at the arithmetic mean, except by invoking response bias. This is also true of the versions of SET, which use a ratio decision rule. This is why Wearden (1991) introduced his version of SET, which uses a difference decision rule. This version predicts a bisection point at the arithmetic mean of the anchor durations.

Hence, at this point, Wearden's (1991) variant of SET seems to be the best model we have of human performance in a bisection task. If so, then, one predicts that, if Yi's (2009) SDT analysis were replicated with human participants, the results would be very different from the one obtained in rats. Instead of remaining constant with stimulus duration, we should find that the standard deviation of subjective time increases linearly with the duration represented. Moreover, the mean of subjective time should also be a linear function the duration represented. We aim to test these two predictions in this study.

Method

In this study, participants are exposed to a 1-s vs. 1.5-s bisection procedure. After being shown a 1-s stimulus and taught to categorize such stimulus as "short" and a 1.5-s stimulus and taught to categorize such stimulus as "long", they are exposed to trials where the actual stimulus duration will range from 1s to 1.5 s.

In order to be able to retrieve the parameters of the SDT model for each participant, the decision criterion λ has to be manipulated (see Wickens, 2002 for the basic theory). Though this can be done explicitly e.g. by using differential payoffs for each response, it is more often manipulated implicitly by using the confidence rating technique. Participants are asked how confident they are about their choice after categorizing a stimulus. The idea is that each confidence rating corresponds to its own decision criterion. A change in λ can then be simulated through the confidence rating data (see Wickens, 2002 for more details). Hence, in the present study, after having categorized the stimulus as “short” or “long”, the participant will have to indicate how sure he is in his judgment: “not sure”, “sure”, or “very sure”.

Participants and apparatus

42 participants (11 males, 31 females), aged between 18 and 26 years old were recruited for this study. The study had been reviewed and approved by the local ethical committee. The experiment was conducted on a PC equipped with an AZERTY keyboard, in a room at the laboratory of affective and cognitive sciences (SCALab) of the university of Lille. A custom MATLAB program controlled the whole procedure.

Procedure

Training. The participant initially read instructions telling them that they would be shown a short-duration and a long-duration stimulus that they will have to categorize as either “short” or “long” (by pressing the left “Ctrl” key for half of the participants, the right “Ctrl” key for the other half) before providing a confidence rating (“not sure”, “sure”, “very sure”) for their categorization.

The experiment proper then started. Sitting 60 cm from the screen, the participant was shown an instruction screen, appearing on white background, informing him that, as soon as he presses a key he would be presented with the short stimuli 4 times. This was followed, after the participant pressed a key, by 4 presentations of 1-s purple rectangle appearing in the center of the screen on a dark background. Each stimulus presentation was separated from the other by a 1-s dark screen. Another instruction screen then informed the participant that he would be presented with the long-duration stimulus 4 times once he pressed a key on the keyboard. Once he did, he was presented with the same purple rectangle appearing 4 times 1 s apart. The only difference was that it was now lasting 1.5 s.

Testing I. The first testing phase started immediately once the training phase was completed. After being prompted by an instruction screen, the participant was exposed to 300 trials. A trial was composed of 3 phases: (a) The stimulus presentation phase: Following a 1-s dark screen, the purple square was shown in the center of the screen for either 1, 1.1, 1.2, 1.3, 1.4 or 1.5 s. Each stimulus duration was shown 50 times. The order of presentation of the stimulus duration was determined randomly. (b) The duration judgment phase: The question "Was this the short-duration or the long-duration stimulus?" appeared on the screen. 22 of the participant answered "short" by pressing the left "Ctrl" key and "long" by pressing the right "Ctrl" key. This was reverse for the remaining participants. Pressing the key caused the question to disappear and the next phase in the trial to start. (c) The confidence judgment phase: The question "How sure are you in your answer?" appeared on the screen with the three possible answers appearing below: "Not sure" on the left, "Sure" in the center, and "Very sure" on the right. Below each answer, the participant was reminded of which key to press in order to select it. The participant chose "Not sure" by pressing the "1" key located the upper left corner of the AZERTY keyboard; "Sure" by pressing the "6" key located in the upper center part of the AZERTY keyboard and "Very Sure" by pressing the "0" key located in the upper right corner of the AZERTY keyboard.

Testing II. For the convenience of the participant, a break took place mid-session to allow him to relax and rest while staying inside the room where the study took place. Once he was ready, the second testing phase began. It was identical to the first one. As they were separated by just a few minutes, they are treated as a single 600-trial session as far as data analysis is concerned.

Data analysis

The data analysis followed steps similar the one used by Maia, Lefèvre, & Jozefowiez (in press). For each participant, the 17 parameters of the SDT model (the mean and standard deviation of the six subjective time distributions corresponding to the six stimulus duration plus the Yes/No response threshold and the four confidence rating criteria) were retrieved from the confidence rating data using an iterative maximum-likelihood algorithm (MLA), as recommended by Wickens (2002) and Macmillan & Creelman (2005). We used the MLA implemented by Harvey's (2013) RscorePlus program. The program "uses singular value decomposition, combined with a variation of the Marquardt method for nonlinear least-squares regression (Marquardt, 1963; Press, Teukolsky, Vetterling, & Flannery, 2002, 2007), to find the maximum likelihood fit of the multiple distribution, a variable-criterion signal detection model to confidence rating-scale data" (Harvey, 2013). Parameters of an SDT model are always expressed relative to a reference condition the mean of which is set to 0 and standard deviation to 1. We chose the 1-s stimulus duration to be that reference. As ratings with a frequency of 0 could prevent the MLA from converging on a solution, a loglinear correction was used, where 1/m (m being the number of possible responses given by the participant, in our case, m = 6) is added to all the data (See Hautus & Lee, 1998).

Based on the likelihood computed by the program that the data could have been generated by the SDT model, a χ^2 (with 15 degrees of freedom) was computed to quantify the fit between the model and the data. A significant χ^2 at the conventional threshold of 0.05 was interpreted as indicating a potential discrepancy between the model and the data.

In the following, we focus on the estimates of the mean and standard deviation of the subjective contingency. We do not report the decision and confidence thresholds as they were of no particular interest to us but we used λ, the criterion which marks the border between "short" and "long" response to compute one of the conventional measure of response bias in SDT. Let's assume that, for a given participant, the SDT analysis has found that when presented with a 1.5-s stimulus, the subjective stimulus duration is drawn from a Gaussian distribution with mean μ and standard deviation σ. By convention, when presented with 1-s stimulus, the subjective stimulus duration is drawn from a standard Gaussian distribution (mean 0 and standard deviation 1). Then, it can be shown that if hits (correctly identifying a 1.5-s stimulus) and false alarms (wrongly identifying a 1-s stimulus) have the same payoff and frequency, the optimal location for the criterion λ is be at the point where the two Gaussian densities cross (Wickens, 2002). If we note β the likelihood ratio of the two Gaussian densities at the criterion, then, if the criterion is placed optimally, β should be equal to 1. That's why a conventional measure of response bias is the logarithm of the likelihood ratio of the densities at the criterion

$$\ln\beta = \ln\left[\frac{F(\lambda,\mu,\sigma)}{F(\lambda,0,1)}\right] = \ln F(\lambda,\mu,\sigma) - \ln F(\lambda,0,1)$$

where $F(x,m,d)$ is the equation for the Gaussian density with mean *m* and standard deviation *d*. $\ln\beta$ is equal to 0 when the participant is not biased. In the context of our study, it will be below 0 when the participant is biased toward the "long" response (the criterion has shifted left relative to its optimal location, resulting in more "long" responses) and above 0 when the participant is biased toward the "short" response (the criterion has shifted right relative to its optimal location, resulting in more "short" responses).

The analysis below fits a second-order polynomial equation to the function mapping objective stimulus onto either the mean of the subjective stimulus duration distribution or its standard deviation. To allow for a meaningful interpretation of the parameters of the polynomial equation, each objective duration *t* was transformed using the equation $(t - 1.25)/0.25$, creating new recoded values ranging from -1 to 1. These recoded values were used to perform the polynomial regression. This recoding made the quadratic and the linear terms of the polynomial equation independent from each other. In these conditions, the linear term corresponds to the slope of the least-mean square linear equation that could have been fitted to the data while the quadratic term reflect the deviation from that straight line (Pagès, 2010).

All parameters based on an averaging of individual performance are reported along with their 95% confidence interval (CI), computed using Student's *t* distribution. All numbers are rounded to the nearest second decimal. Averages for the percentage of variance explained by a regression (r^2) and their 95% CI were computed based on the Fisher transform of the individual r^2. The inverse of the Fisher transform was then used to express them in terms of percentage of variance explained.

Results

Fit of the Gaussian SDT model

Psychometric function and bisection point

Parameters of the SDT model: Mean of the subjective duration distribution

Parameters of the SDT model: Standard deviation of the subjective duration distribution

General discussion

References

Allan, L. G. (2002a). Are the referents remembered in temporal bisection? *Learning and Motivation, 33,* 10-31.

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

Allan, L. G., & Gerhardt, K. (2001). Temporal bisection with trial referents. *Perception & Psychophysics*, *63*, 524-540.

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

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

Falmagne, J. C. (2002). *Element of psychophysical theory*. Cambridge, MA: MIT Press.

Gibbon, J. (1981). On the form and location of the Psychometric Bisection Function for time. *Journal of Mathematical Psychology*, *24,* 56-87.

Gibbon, J., & Church, R. M. (1981). Time left: linear versus logarithmic subjective time. *Journal of Experimental Psychology. Animal Behavior Processes*, *7*, 87-107.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of Sciences*, *423*, 52-77.

Harvey, L. O. (2013). *Parameter estimation of signal detection models: RscorePlus user's manual* (Tech. Rep.). Boulder, Colorado 80309-0345 USA http://psych.colorado.edu/~lharvey/Software%20Zip%20Files/RscorePlus_doc_5_8_3.pdf. Department of Psychology and Neuroscience University of Colorado.

Hautus, M. J., & Lee, A. J. (1998). The dispersions of estimates of sensitivity obtained from four psychophysical procedures: Implications for experimental design. *Perception and Psychophysics, 60,* 638-649.

Jozefowiez, J., Staddon, J. E. R., & Cerutti, D. T. (2009). The behavioral economics of choice and interval timing. *Psychological Review, 116,* 519-539.

Jozefowiez, J., & Machado, A. (2013). On the content of learning in interval timing: Representations or associations? *Behavioural Processes*, *95*, 8-17.

Lejeune, H., & Wearden, J. H. (2006). Scalar properties in animal timing : Conformity and violations. *Quarterly Journal of Experimental Psychology*, *59*, 1875-1908.

Macmillian, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge: Cambridge University Press.

Maia, S., Lefèvre, F., & Jozefowiez, J. (in press). Psychophysics of associative learning: Quantitative properties of subjective contingency. *Journal of Experimental Psychology: Animal Learning and Cognition..*

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial & Applied Mathematics, 11,* 431-441.

Pagès, J. (2010). *Statistique générale pour utilisateurs, vol I: Méthodologie.* Rennes, France: Presse Universitaire de Grenoble.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2002). *Numerical recipes in C++: The art of scientific computing (2nd edition).* Cambridge: Cambridge University Press.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical recipes: The art of scientific computing (3rd edition).* New York: Cambridge University Press.

Provasi, J., Rattat, A.-C., & Droit-Volet, S. (2011). Temporal bisection in 4-month-old infants. *Journal of Experimental Psychology. Animal Behavior Processes*, *37*, 108-113.

Stubbs, A. (1968). The discrimination of stimulus duration by pigeons. *Journal of the Experimental Analysis of Behavior*, *3*, 223-238.

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

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

Wearden, J. H., & Lejeune, H. (2008). Scalar properties in human timing: conformity and violations. *Quarterly Journal of Experimental Psychology*, *61*, 569-587.

Wickens, T. D. (2002). *Elementary signal detection theory*. Oxford: Oxford University Press.

Yi, L. (2009). Do rats represent time logarithmically or linearly? *Behavioural Processes*, *81*, 274-279.

Appendix A

Acknowledgements

Jérémie Jozefowiez, Laboratory of Affective and Cognitive Sciences (SCALab UMR CNRS 9193), Université de Lille (Villeneuve d'Ascq, France) & Centro de Investigação em Psicologia (CIPSI), Universidade do Minho (Braga, Portugal); Clément Gaudichon & Francis Mekkass, Laboratory of Affective and Cognitive Sciences (SCALab), Université de Lille (Villeneuve d'Ascq, France); Armando Machado, Centro de Investigação em Psicologia (CIPSI), Universidade do Minho (Braga, Portugal). This research was supported by grant PTDC/PSI-PCO/100861/2008 from the Fundação para a Ciência e a Tecnologia (J. Jozefowiez, principal investigator), by the Portuguese Foundation for Science and Technology and the Portuguese Ministry of Science, Technology and Higher Education through national funding. It was also co-financed by FEDER through COMPETE2020 under the terms of the PT2020 Partnership Agreement (POCI-01-0145-FEDER-007653). We would like to thank Sophie Lepoutre and Charlotte Renaux for their help in running the participants.
Correspondence concerning this article should be addressed to Jérémie Jozefowiez, Laboratory of Affective and Cognitive Sciences (SCALab UMR CNRS 9193), Université de Lille, Campus de Lille3, Domaine Universitaire du Pont de Bois, BP 60149, 58653 Villeneuve d'Ascq Cedex, France. E-mail: jeremie.jozefowiez@univ-lille3.fr.