The Role of Weber's Law in Human Time Perception

Andrew Haigh[1]
Deborah Apthorp[1, 2]
Lewis A. Bizo[1, 3]

[1]School of Psychology, University of New England,
[2]Research School of Computer Science, Australian National University
[3]Faculty of Arts and Social Sciences, University of Technology Sydney

Correspondence to:
Andrew Haigh
Email: ahaigh2@myune.edu.au

Word count: 5336

Abstract

Weber's law predicts that stimulus sensitivity will increase proportionally with increases in stimulus intensity. Does this hold for the stimulus of time – specifically, duration in the milliseconds to seconds range? There is conflicting evidence on the relationship between temporal sensitivity and duration. Weber's law predicts a linear relationship between sensitivity and duration on interval timing tasks, while two alternative models predict a reverse J-shaped and a U-shaped relationship. Based on previous research, we hypothesised that temporal sensitivity in humans would follow a U-shaped function, increasing and then decreasing with increases in duration, and that this model would provide a better statistical fit to the data than the reverse-J or the simple Weber's Law model. In a two-alternative forced-choice interval comparison task, twenty-four participants made duration judgements about six groups of auditory intervals between 100 and 3200 ms. Weber fractions were generated for each group of intervals and plotted against time to generate a function describing sensitivity to the stimulus of duration. Although the sensitivity function was slightly concave, and the model describing a U-shaped function gave the best fit to the data, the increase in the model fit was not sufficient to warrant the extra free parameter in the chosen model. Further analysis demonstrated that Weber's law itself provided the best description of sensitivity to changes in duration.

Keywords: Time Perception, Weber's law, Scalar Property, Human

The Role of Weber's Law in Human Time Perception

The accurate measurement of time is a biological necessity for all organisms, allowing them to align their internal biological cycles with the external cycles on which they depend for survival. Animals use a variety of biological mechanisms to measure time on scales ranging from microseconds to years (Buonomano, 2007). Out of the many ways that biology has found to measure time, the one that is of most direct relevance to behaviour and cognition is that which spans the duration from milliseconds to minutes (Matell & Meck, 2000). This area of timing, often referred to as *interval timing*, is characterised by both a relatively low level of accuracy and a high level of flexibility in measuring intervals on demand (Gibbon et al., 1997).

Given the importance of interval timing, it is surprising to find that its neurobiological mechanisms are still poorly understood (Matell & Meck, 2000). This uncertainty has led to a debate between several rival models describing different mechanisms for timing, each of which predicts a different mathematical relationship between durations of physical time and measures of perceived time (Grondin, 2001).

In time perception research, estimations of duration are treated as measurements of the perceived intensity of the stimulus of time in a way that is analogous to the intensity of any other physical stimulus (Grondin, 2001). As with all types of perception, organisms are unable to perceive variations in duration if those changes fall below a certain threshold, known as the *just noticeable difference* (*JND*). Analysis of the way that these perceptual thresholds change as duration changes can yield useful information about the nature of the processes that an organism uses to measure time (Grondin, 2001). This analysis is informed by models that attempt to predict the relationship between stimulus threshold and stimulus intensity (Grondin, 2010a).

The relationship between threshold and intensity is described by Weber's law (Sowden, 2012). In its strict form, Weber's law predicts that the ratio of *JND* to stimulus intensity (*I*) will be constant,

$$\frac{JND}{I} = k, \qquad (1)$$

where the term *JND*/*I* is known as the *Weber fraction* (*Wf*) and *k* is known as the *Weber constant* (Holway & Pratt, 1936). Weber's law holds for a wide variety of stimuli across a broad range of intensities (Sowden, 2012); however, it is also violated in many instances (Masin, 2009). For example, many types of stimuli exhibit disproportionately low sensitivity to changes in stimulus intensity at very low stimulus intensity (Sowden, 2012). Although Weber's law is not universal, the *Wf* is still widely used as a measure of stimulus sensitivity in models that seek to describe the mechanisms behind experimentally observed variations in perceptual thresholds (Gibbon et al., 1997).

Stimulus thresholds can be derived by asking subjects to discriminate between two stimuli and then plotting the percentage of correct discriminations against stimulus intensity to generate a *psychometric function* (Kingdom & Prins, 2016). By defining the *JND* in terms of the slope of the psychometric function, the variability of the perceptual discriminations becomes a measure of sensitivity to changes in stimulus intensity (Grondin, 2010b). Definition of the JND thus allows the *Wf* to be stated as a coefficient of variation in terms of the ratio between the standard deviation (*SD*) of perceptual discriminations and the mean (*M*) of those discriminations:

$$Wf = \frac{SD}{M} \tag{2}$$

The way that sensitivity to changes in stimulus intensity varies across a specific range can be visualised by plotting the *Wf* against stimulus intensity. The resulting function, the *perceptual sensitivity function* (PSF), can exhibit a variety of shapes depending on the way that the *Wf* varies with changes in stimulus intensity (Lejeune & Wearden, 2006). Weber's law predicts that *Wf*s will be constant across all intensity values, and therefore in situations where Weber's law is supported, the PSF will be flat. In contrast, where stimulus sensitivity is very low at low stimulus intensity but constant at higher intensities, the PSF will have a reverse J shape, with *Wf*s starting high and falling to a horizontal asymptote.

The applicability of Weber's law to the relationship between duration and the perception of duration has been the subject of ongoing debate in the literature (e.g. Bizo et al., 2006; Getty, 1975; Grondin, 2014; Haß et al., 2008; Killeen & Weiss, 1987). *Scalar expectancy theory,* predicts that measurements of stimulus thresholds for time perception will exhibit constant *Wf*s and therefore flat PSFs (Gibbon & Church, 1984), a relationship that has come to be known as the *scalar property* of time perception (Grondin, 2014). This relationship can be stated mathematically by replacing *M* with the mean duration of the temporal discriminations ($\bar{t}$) in Equation 2 (Gibbon 1977).

$$Wf = \frac{SD}{\bar{t}} = k \tag{3}$$

There is good evidence, however, that time perception is not entirely scalar, but violates Weber's law at very short intervals, with *Wf*s that fall from a high value to a horizontal asymptote at a point somewhere between 50 and 2000 ms (Church et al., 1976; Corke et al., 2018; Fetterman & Killeen, 1992; Getty, 1975). Getty (1975) proposed a generalised form of Weber's law which models this characteristic reverse J-shaped PSF according to the equation,

$$Wf = \frac{\sqrt{A\bar{t}^2 + C}}{\bar{t}}, \tag{4}$$

where the parameter *C* represents a component of residual noise variance, *A* is a parameter related to the value of the Weber constant, and *Wf* and $\bar{t}$ are as defined above (for a derivation of this equation, see the Supplementary Materials).

Another prominent model developed to describe the relationship between stimulus sensitivity and stimulus intensity in time perception is that of Killeen and Weiss (1987). This model represents the variability in subjects' ability to measure time in terms of the advantage that they gain from segmenting intervals into subintervals in a way that minimises variance. The result is a quadratic relation in which the PSF is given by,

$$Wf = \frac{\sqrt{A\bar{t}^2 + B\bar{t} + C}}{\bar{t}}, \tag{5}$$

where *A*, *B*, and *C* are free parameters (Killeen & Weiss, 1987; see Supplementary Materials for derivation). The strength of this model is that it accommodates many previously developed models as special cases. For example, Equation 5 becomes Weber's law (Equation 3) with $B = C = 0$, and it becomes Getty's model (Equation 4) with $B = 0$ (Killeen & Weiss, 1987).

The model proposed by Killeen and Weiss (1987) can be used to describe both the flat PSFs and the reverse J-shaped PSFs found in the experimental literature. However, it assumes that the scalar property of time perception is only violated at shorter intervals and holds at longer intervals. The majority of studies in both the human and animal timing literature support this assumption (Lejeune & Wearden, 2006; Wearden & Lejeune, 2008). There is, however, some evidence to suggest that it may not be universally correct, with several studies showing rising *Wf*s at longer intervals (Grondin, 2010b, 2012; Lavoie & Grondin, 2004; Lejeune & Wearden, 1991). Conjoined with earlier findings of falling *Wf*s at shorter intervals, this evidence suggests that the overall shape of the PSF may, at least in some circumstances, be U-shaped rather than reverse J-shaped, falling at shorter intervals only to rise again at longer intervals.

Perhaps the most well-known example of U-shaped PSFs comes from Getty (1975). A two-alternative forced-choice interval comparison paradigm was used to measure temporal perception thresholds in two human subjects at a range of durations from 50 to 3200 ms. *Wf*s were highest at 50 ms and levelled out at around 200 ms; however, careful examination of Getty's data reveals that *Wf*s began to rise again somewhere around 2500 ms (Figure 1). A similar result in a different range of intervals had been reported many years earlier by Woodrow (1930), who measured perceptual thresholds from 0.2 to 30 s using a temporal reproduction task in eight male subjects. Woodrow found *Wf*s that decreased slightly from 0.2 to 0.6 s, remained constant to 1.5 s, and then increased beyond 1.5 s.

U-shaped PSFs are not limited to the human timing literature. Cantor and Wilson (1981) found a U-shaped PSF with low points from 0.5 to 2 s in rats performing a temporal reproduction task across a range of intervals from 0.2 to 6 s. More recently, U-shaped PSFs were found in pigeons using both temporal production and interval comparison paradigms across durations from 0.5 to 64 s (Bizo et al., 2006), and in domestic dogs using a temporal bisection paradigm across intervals from 0.5 to 16 s (Cliff et al., 2019).

This small body of experimental evidence for the existence of U-shaped PSFs is problematic. Even the most generalised model of time perception (Killeen & Weiss, 1987) does not accommodate data with *Wf*s that increase at longer intervals. In an attempt to fill this gap, Bizo et al. (2006) modified the Killeen and Weiss (1987) model to describe the U-shaped PSF generated in their study, yielding a *Wf* given by,

$$Wf = \frac{\sqrt{A\bar{t}^m + B\bar{t} + C}}{\bar{t}}, \tag{6}$$

where *m* is an additional free parameter which allows the exponent in Equation 5 to vary its value to fit the data.

Information about the shape of the PSF is important in timing research because it informs the development of models that seek to describe the neurological and cognitive processes that give rise to the perception of time (Grondin, 2010b). Many of these models rely on the assumption that *Wf*s remain constant at longer intervals (Matell & Meck, 2000). However, the research cited above has established that time perception is not always scalar at longer intervals (Grondin, 2010a, 2012; Lavoie & Grondin, 2004; Lejeune & Wearden, 1991). These deviations from the scalar property are made manifest by the U-shaped PSFs that are generated by both humans and animals under some conditions (Bizo et al., 2006; Cantor & Wilson, 1981; Cliff et al., 2019; Getty, 1975; Woodrow, 1930).

Getty's (1975) data, showing that human PSFs rise after about 2500 ms (Figure 1), has been cited as evidence of the violation of the scalar property of time perception at longer intervals (e.g. Bizo et al., 2006; Grondin, 2001, 2010; Haß et al., 2008; Lavoie & Grondin, 2004). The results of human studies by Woodrow (1930), Grondin (2010a), Grondin (2012), and Lavoie and Grondin (2004) suggest that this rise might begin as early as 1200 ms. Other studies have failed to find rising PSFs in human subjects (Wearden & Lejeune, 2008). Is this rise a reliable effect? And if so, at what point in the overall range of millisecond to minutes scale timing does it occur?

We aimed to explore the anomaly in Getty's (1975) data to determine whether variations in temporal sensitivity in humans are best described by a reverse J-shaped or U-shaped PSF. We used a methodology similar to that utilised by Getty to generate *Wf*s across a range of intervals from 100 to 3200 ms. The resulting data was then fit to the generalised model of Killeen and Weiss (1987; Equation 5) and the variant of that model developed by Bizo et al. (2006; Equation 6). We hypothesised that the PSF generated from this dataset would be U-shaped. We also hypothesised that the best fit for this function would be given by the model developed by Bizo et al., which is the only extant model capable of describing U-shaped PSFs.

Method

Participants

The sample consisted of 24 participants, 14 of whom were female (58%). Participants ranged in age from 24 to 73 years (*M* = 38.13, *SD* = 10.18), and had adequate hearing for the experimental task, and were able to understand written instructions in English.

Participants were recruited via an invitation circulated through the online social media platform Facebook and gave consent via an electronic form presented at the beginning of the experiment. The Human Research Ethics Committee of the University of New England approved this study (HE19-075).

Apparatus and Materials

The experiment was carried out using a custom-made script running on version 5.0 of the Inquisit software platform (Millisecond, 2018). The scrip is available in the Supplementary Materials. The stimuli were defined by the start and stop points of a series of pure 440 Hz tones of different durations. These tones were generated using version 8.5 of the professional digital audio workstation Cubase (Steinberg, 2015) and recorded as WAV files (available in the supplementary files).

The stimuli consisted of two types of intervals, standard intervals (SI) and test intervals (TI), which were identical in all aspects apart from their duration. The six experimental conditions (S1 to S6) were defined by the six SIs, which were distributed in logarithmic increments from 100 to 3200 ms. For each SI there were 5 TIs, consisting of a central TI equal to the duration of the SI itself and two TIs either side of the SI spaced at durations proportional to the magnitude of the SI (Table 1).

The experiment was run on a Lenovo Yoga 520 laptop, and the audio stimuli were delivered binaurally using standard Audio-Technica ATH-M20x headphones at an A-weighted sound level of 55 dB. Responses were indicated using the left and right arrow keys on the computer keyboard.

Procedure

Participants were tested individually in single sessions lasting between 60 and 90 minutes. The experiment was conducted in a small enclosed office with participants seated facing away from the windows to minimise the risk of distraction. The A-weighted ambient sound level in the room was 32.6 dB, and the light intensity was 411 lx.

The experiment utilised a two-alternative forced-choice interval comparison task similar to that used by Getty (1975). In each trial, a single pair of stimuli consisting of an SI and a TI were presented. Participants were required to decide which of the two intervals was longer. The experiment consisted of six blocks of 100 trials each (Figure 2, Panel A). Each experimental block tested one of the six SIs — each corresponding to one of the six experimental conditions. The order of the blocks was permuted using a balanced Latin square to minimise the risk of order effects between the six conditions. Each block consisted of 20 randomly distributed comparisons between each of the five TIs and the SI of that condition (Figure 2, Panel B). The order of presentation of the SI and TI was randomly varied to avoid interval order effects (Jamieson & Petrusic, 1975). Interstimulus intervals were varied randomly from 750 to 850 ms, and the post-stimulus interval was varied randomly from 350 to 450 ms so that there were no regular intervals in the experiment apart from the experimental stimuli.

Each experimental session was initiated with two short training blocks of 20 trials each. In the first training block, which was designed to test participants' understanding of the instructions, the intervals were easily distinguishable, and feedback was given after each response. In the second training block, the intervals were identical to those of the third (S3) condition, and there was no feedback following each response by a participant. During these two training blocks, the experimenter watched from outside the room to be available to answer any questions. Participants were then left to complete the remainder of the experiment. All participants performed satisfactorily on the first block of practice trials and were able to complete the experiment without assistance.

Participants were instructed at the beginning of each block not to tap, count, or use any other periodic movements to measure the intervals. This was done to reduce the likelihood that participants would use counting as a mediating strategy. The level of participants' confidence in their ability to comply with this instruction was assessed at the end of each block using a five-point Likert scale between "very uncertain" and "very certain".

At the end of each block, participants were able to take a break for whatever duration they desired, and refreshments were available in the experiment room throughout the experiment. Participants were debriefed and allowed to ask additional questions about the experiment at the end of the experiment.

Data analysis

Initial data analysis was conducted using Microsoft Excel (Microsoft Office 365 Pro Plus, Version 16.0) and its Solver addons. Three metrics for response accuracy, practice effects, and response bias were calculated to assess the validity of the data. Response accuracy was assessed by recording a value of 1 for correct responses and 0 for incorrect responses in each trial. Because the third TI was identical to the SI and therefore neither response could be correct or incorrect, responses to this TI were assigned a value of 0.5 in this accuracy metric regardless of the judgement made. This ensured that random responding would result in 50% accuracy in this metric (Note that this data was not used for calculating the *Wf*). Practice effects were assessed by creating a separate practice metric and assigning a value of 1 to each correct judgement, excluding the third TI. These values were totalled over the entire experiment to generate a cumulative number correct, which was plotted against the trial number and assessed for linearity using least-squares linear regression. Response biases were assessed using the response values recorded in the raw data to generate a value for the percentage of right-arrow responses in each block.

The raw data used to generate the *Wf*s consisted of 20 binary discriminations (SI longer or TI longer) for each of the 30 TIs. Discriminations were given a value of 0 if the participant judged the SI to be the longer of each pair of intervals and a value of 1 if the participant judged the TI to be longer. These values were summed across the 20 trials of each of the five TIs in each block. They were then converted into a percentage, yielding a value for the percentage of long responses for each TI. These percentage long values were plotted against TI duration within each block to give six psychometric functions for each participant. The standard deviation for each of these psychometric functions was calculated directly from a frequency distribution consisting of the time intervals for which “longer” judgements were made in each block, and the *Wf* for each condition was calculated by dividing the standard deviation by the mean discrimination duration ($\bar{t}$).

The independent variable for our primary analysis was the duration of the SIs in the six conditions, while the dependent variable was the *Wf* generated in each condition. To assess the degree of difference between these six *Wf*s, a repeated-measures analysis of variance was conducted using the SPSS software package (IBM SPSS Statistics, Version 23). PSFs for each participant were generated by plotting the *Wf* against $\bar{t}$ for each condition. An overall mean PSF was generated by plotting the mean of all participant’s *Wf*s against overall mean $\bar{t}$ for each condition (see Figure 4).

A coefficient of determination (r^2) describing the degree of fit between these PSFs and the Killeen and Weiss (1987; Equation 5) and Bizo et al. (2006; Equation 6) models was calculated according to the non-linear regression procedure outlined by Brown (2001) using the Solver plug-in in Microsoft Excel (Microsoft Office 365 Pro Plus, Version 16.0). The difference in the number of free parameters in the two models made interpretation of coefficients of determination problematic (Equations 5 and 6 with three and four free parameters, respectively; Spiess & Neumeyer, 2010). Consequently, the final analysis was conducted using corrected Akaike information criterion (AICc) values (Burnham & Anderson, 2004; for mathematical details, see the Supplementary Materials).

Results

Mean discrimination durations ($\bar{t}$) were consistently close to the SI duration (Supplementary Materials, Figure S1), and all mean psychometric functions had a positive slope (Figure 3). The six psychometric functions for each of the 24 participants (see Supplementary Materials, Figure S2) also had a positive slope, apart from two cases with zero slopes and one with a slightly negative slope (participant 1, S1 condition; Figure S2). Replacing the latter *Wf* with a value corresponding to a zero slope did not affect the outcome, so all psychometric functions were included in the data analysis.

*Wf*s ranged from 0.05 to 0.12 (Supplementary Materials, Table S2), and mean *Wf*s were between 0.09 and 0.10 (M = 0.09, SD = 0.01; Table 2). The overall mean PSF was slightly concave (Figure 4, black dots). A one-way repeated-measures analysis of variance showed that there was a significant difference in the *Wf*s between the six conditions, $F(5, 115) = 4.69$, p = .001, with a large effect size (partial $\eta^2 = .17$). Pairwise comparisons using the Bonferroni correction revealed that the *Wf*s for the shortest condition (S1) were significantly higher than those of the third, fourth, and fifth conditions ($p < .05$; Table 3). Although the mean *Wf* for the longest (S6) condition was also higher than that of the third, fourth, and fifth conditions (Table 2), the difference was not statistically significant, and no other comparisons were statistically significant (Table 3).

The Bizo et al. (2006) model had a higher coefficient of determination when fit to the overall PSF (r^2 = 0.80) than the Killeen and Weiss (1987) model (r^2 = 0.60). The Bizo et al. model also had a higher mean coefficient of determination when fit to the individual PSFs (mean r^2 = 0.46, SD = 0.28) than the Killeen and Weiss model (mean r^2 = 0.26, SD = 0.30; Table 4; for individual PSFs with 95% confidence intervals see Supplementary Materials, Figure S2). The Bizo et al. model gave a better fit to the data than the Killeen and Weiss model when no adjustments are made for the difference in the number of free parameters between the two models. There was a high degree of variation in the coefficient of determination values, ranging from 0 to 0.95 for the Bizo et al. model and -0.34 to 0.95 for the Killeen and Weiss model (Figure 5).

AICc values for the fit to the overall PSF were higher for the Bizo et al. model than for the Killeen and Weiss model (Table 4, upper portion), with an AIC_c difference value between the two models of 28.21. The mean AICc values for the fit to the individual PSFs were also higher for the Bizo et al. model than for the Killeen and Weiss model (Table 4, lower portion) with an AIC_c difference value between the two models of 29.09. This demonstrates that the Killeen and Weiss model gave the best fit to the data when adjusted for the difference in the number of free parameters between the two models (Burnham & Anderson, 2004).

The mean percentage correct for each condition was consistently above 60% (M = 69%, SD = 5.46; Figure 6). A one-sample t-test comparing the percentage correct with the 50% value expected with random responding found that mean percent correct differed significantly from chance in all six conditions ($p < .001$). Plots of the cumulative number of trials with correct responses across the experiment for each participant (see Supplementary Materials, Figure S3) showed an almost perfectly linear accuracy pattern for all participants (mean $r^2 = 1.00$), indicating that there were no learning effects in this experiment.

The results of the assessment of response bias were mixed. The overall percentage of right-arrow responses was 50% (*SD* = 6.04), which is the value expected from a bias-free response pattern; however, right-arrow responses were at their lowest number in the shortest interval condition, *M* = 39.83, *SD* = 12.69, and climbed steadily to their highest number in the longest interval condition, *M* = 63.04, *SD* = 10.02 (Figure 7). A one-sample *t*-test comparing the percentage of right-arrow responses with a bias-free performance of 50% showed that the bias was significant in both the S1 condition, $t(23) = - 3.93$, $p = .001$, $d = 0.80$ and the S6 condition, $t(23) = 6.37$, $p < .001$, $d = 1.30$.

Participants tended to report a high level of confidence that they were not counting or using any other rhythmic strategies in this experiment (*M* = 4.0, *SD* = 1.18). Mean confidence levels were highest in the S1 condition (*M* = 4.54, *SD* = 0.98). and declined to their lowest level in the S6 condition (*M* = 3.58, *SD* = 1.44; Figure 8). A Pearson's correlation between participants' coefficients of determination in the fit to the Bizo et al. (2006) model and their responses to the confidence question for the longest (S6) interval condition revealed a weak correlation that was not statistically significant, $r(24) = .18$, $p = .396$.

Discussion

Overall, participants in this study were more likely to judge the TI as being longer than the SI as TI duration increased (Supplementary Materials, Figure S1), which demonstrates that the stimuli were within the intended range of participants' sensitivity to duration. Participants also exhibited high accuracy in their temporal discriminations, with consistently low *Wf*s (see Supplementary Materials, Table S2) falling within a range similar to that found in Getty's (1975) original study (Figure 1).

The hypothesis that the PSF would be U-shaped was only weakly supported. Visually, the mean PSF had a slightly concave shape (Figure 4), and mean *Wf*s were marginally higher in the two shortest (S1 and S2) and the longest (S6) conditions (Table 2); however, the slight upturn in the longest interval, which is the crucial element in demonstrating a U-shaped PSF, was not statistically significant (Table 3). Furthermore, the PSFs of individual participants did not show any consistent pattern (see Supplementary Materials, Figure S2), whereas the PSFs of both subjects in Getty's (1975) study are U-shaped (Figure 1). Visual inspection of Figure 1 demonstrates that the drop in *Wf*s at short intervals found in this study is smaller than that found by Getty; however, part of this early drop in *Wf*s reported by Getty occurs between the 50 and 100 ms intervals, whereas the 50 ms interval was not included in this study. Ignoring the first data point in both panels of Figure 1, the main difference between these two functions and the mean PSF found in this study (Figure 4) is the markedly lower magnitude of the rise in *Wf*s at longer intervals in the latter compared with the former.

The hypothesis that the Bizo et al. (2006) model (Equation 6) would give the best fit to the data was also only weakly supported. When using the coefficient of determination as the metric of comparison, the Bizo et al. model gave a better fit to both the mean PSF and the individual PSFs. This result, however, did not hold for the comparison of the AICc values, which adjust for the different number of free parameters in the two models (Burnham & Anderson, 2004). The magnitude of the AIC_c differences between these two models for both the mean PSF and the individual PSFs are high enough to conclude that the hypothesis that the Bizo et al. model would give a better fit to the data has no empirical support (Burnham et al., 2011). Therefore, although allowing the exponent of the first term in the Killeen and Weiss (1987) model (Equation 5) to vary (Equation 6) did yield a higher coefficient of determination for the Bizo et al. model, the increase in the accuracy of the model fit to this dataset was not sufficient to justify the addition of an extra free parameter into the model.

Given that both Getty's (1975) model (Equation 4) and Weber's law itself (Equation 3) have fewer free parameters than either the Bizo et al. (2006) or the Killeen and Weiss (1987) models, it is useful to explore how the former two models compare to the latter in the model fit to this dataset. A comparison between all four models (Table 4) shows that, although the Bizo et al. model certainly had the highest coefficients of determination of the four, it is Weber's law, the model with the least number of free parameters, which has the lowest AICc value. This demonstrates that Equation 3 provides the best fit to this dataset when the differences in the number of free parameters is taken into account. Thus, the temporal sensitivity of the participants in this study is best described by a linear function as Weber's law predicts (Holway & Pratt, 1936).

There are, however, a few considerations that must qualify any generalisations based on these results. This study sought to examine the general trend in timing accuracy across a specific range of intervals. It did not attempt to address the question of how counting affects timing accuracy across that range. There is conflicting evidence on the effect of counting on the accuracy of temporal judgements in humans (Hinton & Rao, 2004), but it is generally assumed that counting improves accuracy at longer intervals, therefore lowering *Wf*s (Fetterman & Killeen, 1990). Although participants were encouraged not to count in both the current study and Getty's original (1975) study, the possibility remains that the difference in the profile of the PSFs between the two studies is the result of a higher level of motivation to comply with this directive in Getty's participants. This possibility is supported by the results of the confidence question, which show that participants in this study were the least confident that they refrained from counting in the longer intervals (Figure 8). There was no evidence for a systematic relationship between participant's perceptions of their ability to resist counting and the fit with the Bizo et al. model, however. To further explore this relationship, subsequent research could look specifically at the effect of counting on timing performance.

The results of this study do not seem to have been affected by random responding or learning effects. There was a significant systematic bias in the responses; however. Left-arrow responses predominated in the S1 condition right-arrow responses predominated in the S6 condition. Due to the randomisation in the order of presentation of SIs and TIs, this bias is not indicative of an interval order effect. This might reflect an inherent bias to associate shorter intervals with the left arrow and longer intervals with the right arrow; however, further research would be required to elucidate the nature of this effect.

This study was conducted with a relatively large sample size which had a good spread of ages and a reasonable gender balance. The results of this study suggest that the decrease in accuracy of temporal discriminations at longer intervals found by Getty (1975) is not a generalisable effect. In addition, our results and Getty's results together suggest that different individuals may have different profiles of sensitivity to changes in duration. Further research using a similar procedure on a larger sample of participants could establish whether there are indeed significant individual differences in the profile of the PSF for time, and if so, what the behavioural and cognitive correlates of these differences might be.

Individual differences in sensitivity to time are known to exist. For example, deficits in time perception have been found in a range of psychological and neurological conditions (Gibbon et al., 1997), including Parkinson's disease (Malapani, Rakitin, Levy, & Meck, 1998), Alzheimer's disease (Haj & Kapogiannis, 2016), and Schizophrenia (Ueda, Maruo, & Sumiyoshi, 2018). Because the paradigm used in this experiment was designed to run on a standard commercially available software platform, the current study provides a reproducible procedure which could be used to explore variations in the profile of temporal sensitivity across the human population.

Numerous variants of Weber's law have been proposed to model the profile of sensitivity to changes in the physical stimulus of duration, two of which were compared in this study. Although the most sophisticated of these models (Bizo et al., 2006) provided the best raw fit to the data, the increase in the fit between the two models was not sufficient to warrant the extra free parameter required. Furthermore, accommodations for the difference in the number of free parameters revealed that the model with the smallest number of free parameters, Weber's law itself, actually gave the best fit to the data. This result demonstrates that the decrease in sensitivity to the stimulus of duration found in some previous research at intervals between 1 and 3 s is not a consistent effect. It also adds to the large body of evidence demonstrating that, in certain situations and within a certain range of intervals, the profile of sensitivity to changes in duration is best described by Weber's law. Thus, at least in the case of this research, Weber's law appears to have stood the test of time.

References

Bizo, L. A., Chu, J. Y. M., Sanabria, F., & Killeen, P. R. (2006). The failure of Weber's law in time perception and production. *Behavioural Processes*, *71*(2–3), 201–210. https://doi.org/10.1016/j.beproc.2005.11.006

Brown, A. M. (2001). A step-by-step guide to non-linear regression analysis of experimental data using a Microsoft Excel spreadsheet. *Computer Methods and Programs in Biomedicine*, *65*(3), 191–200. https://doi.org/10.1016/S0169-2607(00)00124-3

Buonomano, D. V. (2007). The biology of time across different scales. *Nature Chemical Biology*, *3*(10), 594–597. https://doi.org/10.1038/nchembio1007-594

Burnham, K. P., & Anderson, D. R. (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Burnham, K. P., Anderson, D. R., & Huyvaert, K. P. (2011). AIC model selection and multimodel inference in behavioral ecology: Some background, observations, and comparisons. *Behavioral Ecology and Sociobiology*, *65*(1), 23–35. https://doi.org/10.1007/s00265-010-1029-6

Cantor, M. B., & Wilson, J. F. (1981). Temporal uncertainty as an associative metric: Operant simulations of Pavlovian conditioning. *Journal of Experimental Psychology: General*, *110*(2), 232–268. https://doi.org/10.1037/0096-3445.110.2.232

Church, R. M., Getty, D. J., & Lerner, N. D. (1976). Duration discrimination by rats. *Journal of Experimental Psychology: Animal Behavior Processes*, *2*(4), 303–312. https://doi.org/10.1037/0097-7403.2.4.303

Cliff, J., Jackson, S., McEwan, J., & Bizo, L. (2019). Weber's law and the scalar property of timing: A test of canine timing. *Animals, 9*(10), 801. https://doi.org/10.3390/ani9100801

Corke, M., Bell, J., Goodhew, S. C., Smithson, M., & Edwards, M. (2018). Perceived time slows during fleeting fun or fear. *Quarterly Journal of Experimental Psychology*, *71*(2), 562–567. https://doi.org/10.1080/17470218.2016.1264000

Fetterman, J. G., & Killeen, P. R. (1990). A componential analysis of pacemaker-counter timing systems. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(4), 766–780. https://doi.org/10.1037/0096-1523.16.4.766

Fetterman, J. G., & Killeen, P. R. (1992). Time discrimination in Columba livia and Homo sapiens. *Journal of Experimental Psychology: Animal Behavior Processes*, *18*(1), 80–94. https://doi.org/10.1037/0097-7403.18.1.80

Getty, D. J. (1975). Discrimination of short temporal intervals: A comparison of two models. *Perception & Psychophysics*, *18*(1), 1–8. https://doi.org/10.3758/BF03199358

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, *84*(3), 279–325. https://doi.org/10.1037/0033-295X.84.3.279

Gibbon, J., & Church, R. M. (1984). Sources of variance in an information processing theory of timing. In H. Roitblat, H. Terrace, & W. Sever (Eds.), *Animal cognition* (pp. 465–488). Erlbaum.

Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. R. (1997). Toward a neurobiology of temporal cognition: Advances and challenges. *Current Opinion in Neurobiology*, *7*(2), 170–184. https://doi.org/10.1016/S0959-4388(97)80005-0

Grondin, S. (2001). From physical time to the first and second moments of psychological time. *Psychological Bulletin*, *127*(1), 22–44. https://doi.org/10.1037/0033-2909.127.1.22

Grondin, S. (2010a). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics*, *72*(3), 561–582. https://doi.org/10.3758/APP.72.3.561

Grondin, S. (2010b). Unequal Weber fractions for the categorization of brief temporal intervals. *Attention, Perception & Psychophysics*, *72*(5), 1422–1430. https://doi.org/10.3758/APP.72.5.1422

Grondin, S. (2012). Violation of the scalar property for time perception between 1 and 2 seconds: Evidence from interval discrimination, reproduction, and categorization. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 880–890. https://doi.org/10.1037/a0027188

Grondin, S. (2014). About the (Non)scalar Property for Time Perception. In H. Merchant & V. de Lafuente (Eds.), *Neurobiology of Interval Timing* (Vol. 829, pp. 17–32). Springer New York. https://doi.org/10.1007/978-1-4939-1782-2_2

Haß, J., Blaschke, S., Rammsayer, T., & Herrmann, J. M. (2008). A neurocomputational model for optimal temporal processing. *Journal of Computational Neuroscience*, *25*(3), 449–464. https://doi.org/10.1007/s10827-008-0088-4

Hinton, S. C., & Rao, S. M. (2004). “One-thousand one … one-thousand two …”: Chronometric counting violates the scalar property in interval timing. *Psychonomic Bulletin & Review*, *11*(1), 24–30. https://doi.org/10.3758/BF03206456

Holway, A. H., & Pratt, C. C. (1936). The Weber ratio for intensive discrimination. *Psychological Review*, *43*(4), 322–340. https://doi.org/10.1037/h0059748

Killeen, P. R., & Weiss, N. A. (1987). Optimal timing and the Weber function. *Psychological Review*, *94*(4), 455–468. https://doi.org/10.1037/0033-295X.94.4.455

Kingdom, F., & Prins, N. (2016). *Psychophysics: A Practical Introduction* (2nd ed.). Academic Press. https://doi.org/10.1016/B978-0-12-407156-8.12001-8

Lavoie, P., & Grondin, S. (2004). Information processing limitations as revealed by temporal discrimination. *Brain and Cognition*, *54*(3), 198–200. https://doi.org/10.1016/j.bandc.2004.02.039

Lejeune, H., & Wearden, J. H. (1991). The comparative psychology of fixed-interval responding: Some quantitative analyses. *Learning and Motivation*, *22*(1–2), 84–111. https://doi.org/10.1016/0023-9690(91)90018-4

Lejeune, H., & Wearden, J. H. (2006). Scalar properties in animal timing: Conformity and violations. *Quarterly Journal of Experimental Psychology*, *59*(11), 1875–1908. https://doi.org/10.1080/17470210600784649

Masin, S. (2009). The (Weber's) law that never was. In M. Elliott & S. Antonijevic (Eds.), *Fechner Day 2009: Proceedings of the 25th Annual Meeting of the International Society for Psychophysics* (pp. 441–446). International Society for Psychophysics.

Matell, M. S., & Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology*, *22*(1), 94–103. https://doi.org/10.1002/(SICI)1521-1878(200001)22:1<94::AID-BIES14>3.0.CO;2-E

Millisecond. (2018). *Inquisit* (Version 5.0) [Computer software]. Millisecond Software. https://www.millisecond.com

Sowden, P. T. (2012). Psychophysics. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol 1: Foundations, planning, measures, and psychometrics.* (pp. 445–458). American Psychological Association. https://doi.org/10.1037/13619-024

Spiess, A.-N., & Neumeyer, N. (2010). An evaluation of R2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: A Monte Carlo approach. *BMC Pharmacology*, *10*(1), 6. https://doi.org/10.1186/1471-2210-10-6

Steinberg. (2015). *Cubase* (Version 8.5) [Computer software]. Steinberg Media Technologies.

Wearden, J. H., & Lejeune, H. (2008). Scalar properties in human timing: Conformity and violations. *Quarterly Journal of Experimental Psychology*, *61*(4), 569–587. https://doi.org/10.1080/17470210701282576

Woodrow, H. (1930). The reproduction of temporal intervals. *Journal of Experimental Psychology*, *13*(6), 473–499. https://doi.org/10.1037/h0070462