A Test of Weber's Law

with Dogs

A thesis

submitted in partial fulfilment

of the requirements for the degree

of

Master of Applied Psychology (Behaviour Analysis)

at

The University of Waikato

by

JESSICA HELEN CLIFF

THE UNIVERSITY OF
WAIKATO
Te Whare Wānanga o Waikato

2013

Abstract

This study investigated Weber's Law through a temporal bisection procedure using domestic dogs as subjects, a previously untested species in temporal bisections. Six dogs were trained to response to the blue lever when the short signal duration was presented, and response to the red lever when the long signal duration was presented. The four conditions were 0.5 – 2.0 s, 1.0 – 4.0 s, 2.0 – 8.0 s and 4.0 – 16.0 s. The intermediate durations presented were logarithmic intervals of the two original signal durations. On each trial there were 7 possible durations for the green light, the two trained durations and the 5 intermediate logarithmic durations between them. Reinforcement was provided for correct responses to trained durations through out training and testing, no reinforcement was given for intermediate durations.

This study demonstrated the PSE was close to the geometric mean and a failure of Weber's Law. Weber fractions were not constant and instead produced a U-shaped function. Starting large for the shortest condition, (0.5 – 2.0 s), getting smaller for the middle two conditions (1.0 – 4.0 s & 2.0 – 8.0 s) and increased again for the longest condition (4.0 – 16.0 s). Results demonstrated that subjects found discriminating between durations within the 0.5 – 2.0 s condition the hardest. These results replicate the findings of Church and Deluty (1977) where the PSE was also close to the geometric mean, as well as Bizo et al. (2006) and Zeiler (1991) who were both able to demonstrate a failure of Weber's Law.

Acknowledgements

I would like to gratefully and sincerely thank my supervisors Dr. Lewis A. Bizo and Dr. James S. McEwan for their support and guidance throughout this process. It has been an adventure. Thanks to Rob Bakker for his technical assistance with the programming of my experiment and his ingenuity in thinking up and building the dog's apparatus.

I would like to acknowledge the owners of the dogs not only for their dedication to drop offs and pick ups but also for getting into the competitive spirit with each other. It was great to have people who were as excited and invested in it as I was. I want to thank the dogs, without you this would not of been possible. I enjoyed every session with each of you and all the bizarre behavior you brought to each day.

Finally, to my family and friends, thanks for the unwavering support through out my academic career. Thanks especially to my parents, to my mum for walking-and-talking me through my frustrations and my dad for always encouraging me and believing that I can achieve great things. You have both been understanding and supportive every step of the way and without you both I wouldn't be who I am today. Lastly a huge thank you to my husband, your patience and encouragement is never-ending and you show me everyday how important it is to work hard, without you I would never of made it to this point.

Table of Contents

List of Tables

List of Figures

List of Appendices

Introduction

Timing is important to animals for the same reasons people find it useful, it enables the anticipation and prediction of events needed for survival and provides the ability to avoid negative aspects of life. Timing is usually described in terms of different timing models, which attempt to describe different underlying processes and structures responsible for timing ability. Popular models include scalar expectancy theory, behavioural theory of timing and learning to time (Machado & Keen, 1999). Each timing model produces individual predictions which when tested either provides support for the model and it's underlying processes or demonstrates flaws in the model. A common task for testing the predictions of a model is a temporal bisection.

Timing abilities with humans, pigeons and mice were demonstrated through an altered FAA (Food anticipatory activity) test, which presents food at regular short interval, followed by a peak interval task (Balsam, Sanchez-Castillo, Taylor, Van Volkinburg, & Ward, 2009). The peak interval task involves two trial types, the first is a fixed interval procedure where reinforcement is given after a response following a tone that signals the end of the fixed interval. The second type of trial is a peak trial, this involves the same stimulus only the interval is 4 times longer than the fixed interval trials and no reinforcement is given. The results found that response rate was at its peak at the fixed interval time and steadily decreased after that time passed. Balsam et al. (2009) concluded that the accuracy and precision of the responses during the peak intervals showed that the subjects were able to time and generalised that this would also be the case for most animals. This claim can be tested through a temporal bisection procedure.

Temporal Bisection

Temporal bisection tasks involve comparing a stimulus duration against two reference durations. A stimulus is shown for a set duration and the participant must indicate whether they considered the duration to be the same as the short or long reference duration. These reference durations are typically the shortest and longest durations that were presented during training trials, and animals are usually trained to reach a specific level of accuracy before adding in intermediate durations for testing, for example Church and Deluty (1977). Church and Deluty's (1977) temporal bisection procedure used four different conditions (1.0 – 4.0, 2.0 – 8.0, 3.0 – 12.0 & 4.0 – 16.0 s). They used the reference durations and equally spaced logarithmic intervals when testing. Figure 1 shows the pattern of responding from rats in their temporal bisection procedure. This psychometric function shows that the shortest signal duration is almost always identified as "short" and when the longest duration is shown it is almost always identified as "long".

A psychometric function also gives an indicator of where the point of bisection is. This is the duration that is selected as being "Long" 50% of the time. The point of bisection can also be called the point of subjective equality (PSE). In most temporal bisection procedures the harmonic, geometric and arithmetic means of the two trained durations are also some of the intermediate duration that are presented in the testing or probing condition.

A Pseudo-Logistic Function (Equation (1)) can be applied to temporal bisection data. This produces a smooth line of best fit to the data points and visually demonstrates the relationship between the probability of responding "Long" and the signal durations presented.

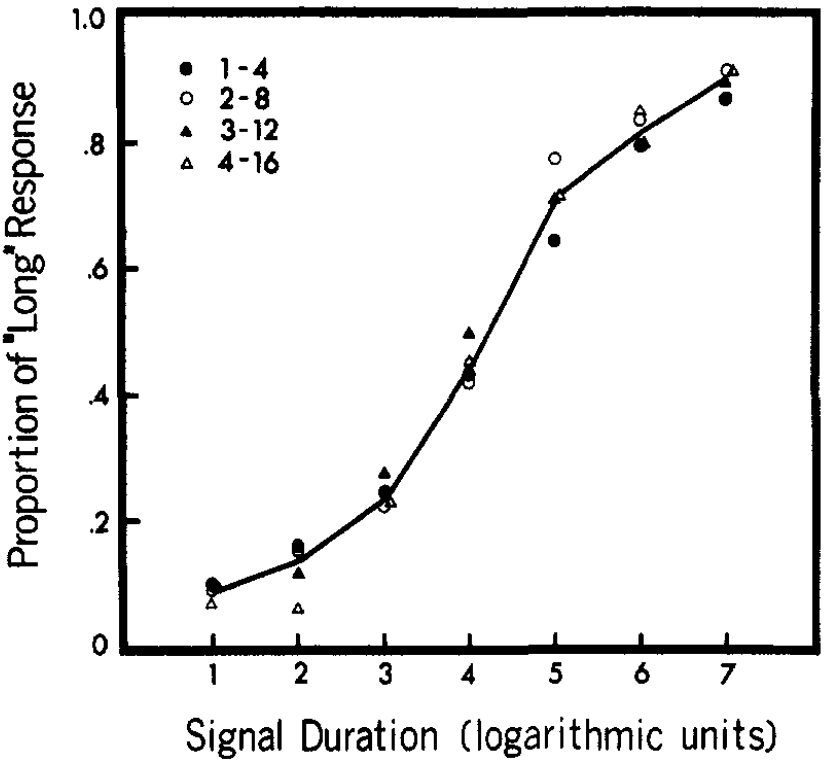


Figure 1. Proportion of long responses as a function of signal duration in logarithmic units. From "Bisection of Temporal Interval" by R. Church and M. Deluty, 1977, *Journal of Experimental psychology: Animal Behavior Processes, 3*, pp. 216-228.

In Equation 1, μ and σ represent the mean and standard deviation, respectively. Equation 1 is suitable for fitting to temporal bisection data as it is not reliant on the proportion of responses following any one alternative becoming exclusive, and reaching 1.0. Something that is unlikely to occur in a temporal bisection procedure (Gibbon, 1977).

$$f = [1 + \exp\left(\frac{\mu - t}{0.55\sigma}\right)]^{-1} \quad (1)$$

Temporal bisection tasks have been conducted with both humans and animals, and some significant differences have been found between the two. The bisection point for animals is closer the geometric mean than the harmonic or arithmetic means, as demonstrated with pigeons (Penney, Gibbon, & Meck, 2008; Stubbs, 1976), rats (Church & Deluty, 1977; Meck, 1983; Raslear, 1983) and mice (Penney et al., 2008). Human studies have produced a difference in the point of bisection being just below the arithmetic mean. In all the conditions of a human temporal bisection study done by Wearden, Rogers and Thomas (1997) the bisection point was closer to the arithmetic mean. In their study with human participants, they found that the shape of the psychometric curve was significantly different depending on whether the stimuli presented were logarithmically spaced or not. Ferrara, Lejeune and Wearden (1997) showed that results can be affected by how far apart the stimuli are spaced. Their study with human participants showed higher sensitivity when stimuli were more closely spaced regardless of whether the spacing was linear or logarithmic. Unlike the findings of Wearden et al. (1997) with human participants, logarithmic spacing does not affect the curve in studies with animal subjects (Church & Deluty, 1977; Wearden et al., 1997).

Wearden et al. (1997) demonstrated that the differences found between human and animal bisection studies were not dependent on the stimuli durations used as Wearden and Ferrara (1995) had previously argued. Wearden et al. (1997) argue that these differences are due to a difference in decision processes not a difference in underlying temporal processes or an internal clock.

Very few temporal bisection procedures have been done with durations under 1.0 s. Zelanti and Droit-Volet (2005) conducted one with people of different ages (5yrs, 9yrs and adults). The signal ranges they used were 0.5 – 1.0 s, 1.25 – 2.5 s, 4.0 – 8.0 s and 15.0 – 30.0 s. They were able to demonstrate a significant difference between the shortest range (0.5 – 1.0 s) compared to the three others but the shorter range also had a different Short/Long (S/L) ratio which could of effected the results.

A temporal bisection procedure has it's advantages over other perception of time tasks as it reduces many of the negative effects that approaches like counting and tapping often have. The impact of a delay in motor response is also reduced for verbal estimations (Lee et al., 2009).

A temporal bisection procedure also provides a basis for experiments comparing relative and absolute durations. For example Church and Deluty (1977) tested which was faster to learn, relative or absolute durations, after the initial temporal bisection procedure, where the rats learned to press the left lever following a signal of 1.0 s in duration and press the right lever following a signal of 4.0 s in duration.

The next temporal discrimination assessed whether the rats learned the discrimination faster if they were in the relative group or the absolute group. Being in the relative group meant that for the new signals of 4.0 s and 16.0 s the short option was still associated with the left lever and the long option still with the right lever, just as it was originally with the short option (1.0 s) on the left lever and the long option (4.0 s) on the right. For the absolute group it was the 4-s signal that had originally been associated with the right lever that continued to be the right lever in the 4 v 16-s discrimination, leaving the 16.0 s duration to be associated with the left lever. This group was called the absolute group as it was the absolute time of 4.0 s that remain consistent across the two experiments. Relative group meant that the shorter option, regardless of its actual length was associated with the left lever in both the initial 1 v 4-s discrimination and the 4 v 16-s discrimination, while being in the absolute group meant that the association between 4.0 s and the right lever stay consistent across the two conditions. Church and Deluty (1977) found that the rats had higher correct proportions if they were in the relative group as opposed to the absolute.

Weber's Law

Weber's Law relates to the intensity or quantity of something, and what is the smallest noticeable difference. Weber's Law states that the just noticeable difference (jnd), the smallest amount of change that is noticeable, is proportional to the original size. For example a tone of 1 s duration is increased to 2 s in length, this is a noticeable difference but if the tone duration was 30 s long a 1 s increase would be unlikely to be noticed without the aid of some device such as a watch. The increase would need to be larger as the original value is larger (Taylor, Haskell, Appleby, & Waran, 2002).

To determine the size of increase needed a Weber fraction is used, this mathematical formula determines the minimum increase that will be noticeable, usually stimuli need to vary by 20% (0.2) for the change to be noticeable. Weber fractions are calculated by dividing the standard deviation by the mean. Church and Deluty (1977) found the average Weber fraction to be 0.23 and there was no significant difference between the four ranges (1.0 – 4.0, 2.0 – 8.0, 3.0 – 12.0 & 4.0 – 16.0 s). A low Weber fraction indicates high discriminability and produces a steep psychometric curve, while a high Weber fraction signals a low ability to discriminate between the testing durations and the psychometric function is flatter and involves a slow gradual rising. For Weber's Law to hold, the Weber fraction should remain constant across varying duration ranges (Lejeune & Wearden, 2006).

Siegel and Church (1984) found evidence with rats that demonstrates Weber fractions decrease as the range between the two trained extreme durations increases. This has been shown in many other temporal bisection studies with a range of animal and human participants including Allan and Gibbon (1991); Droit-Volet and Wearden (2001); and Zeiler (1991) when under a closed economy. Kopec and Brody (2010) suggested that although Weber fractions decrease as the trained duration range increases, it does not violate the Scalar Timing theory of a constant Weber fraction. They looked at the same duration range across difference experiments and found no significant correlation between the arithmetic mean of the trained durations and the Weber fraction.

Kopec and Brody (2010) hypothesized that the reinforcement given during pre-training influences the difference in Weber fractions. This is because with a low range pair of durations the bisection point is close to both extreme trained durations and reinforcement given for that response provides feedback. The long range trained durations result in the participant having no experience in selecting and getting reinforced for selecting a duration close to the bisection point, as only the extreme trained durations are presented in pre-training.

Failure of Weber's Law has been shown in many different timing tasks (Bizo, Chu, Sanabria, & Killeen, 2006; Getty, 1975; Lavioe & Grondin, 2004; Zeiler, 1991; Zelanti & Droit-Volet 2011). Getty (1975) found a failure to predict a rise in the standard deviation for durations 2.0 s or longer and produced a generalised version of Weber's Law which accounted for this finding. Bizo et al. (2006) demonstrated that Getty's generalised Weber's Law still did not account for their findings, as they produced Weber fractions that were U-shaped. Weber's Law and generealised Weber's Law predict flat and J-shaped functions respectively. Despite these findings, Weber's Law and generalised Weber's Law have been used in timing and within some of the most prominent theories. Scalar Expectancy Theory (SET; Gibbon, 1977) relies on Weber's Law and is the most popular timing theory (Malapani & Fairhurst, 2002).

Theories of Timing

There are three popular timing models that attempt to explain and predict the choices humans and animals make regarding judgement of the passage of time. The first is SET which proposes an internal clock, a pacemaker and memory processes are used in decision making, here the number of pulses from the pacemakers are counted and stored in memory. The number of pulses can then be recalled from memory and compared with the current sample to influence expectations and behaviour. Gibbon continued to expand and elaborate SET in later work to account for new findings and information on timing (Gibbon, 1991; Gibbon & Church, 1984).

Killeen and Fetterman (1988) produced an alternative timing model, Behavioural Theory of Timing (BeT), this model includes a pacemaker but suggests that the pacemaker is only used to move the animal through different behaviour states and the animal uses these series of behaviours as signals to occasion responses. Machado (1997) suggests a third alternative to understanding timing in animals, Learning to Time (LeT). This theory is also based around the animals' behaviour and how animals can use their own series of behaviour to measure time but unlike BeT a pacemaker is not used to move from one behaviour state to the next. Response rates during different behaviour states are reflective of the strength of association between those behaviours and reinforcement.

The different predictions that can be assessed through a double bisection task for SET and LeT are presented in Machado and Keen (1999). The double bisection task uses two pairs of trained durations. The first pairing was 1 s red and 4 s green, the second was 4 s blue and 16 s yellow. The different predictions each theory makes are shown in Figure 2, one of the main differences is that while both SET and LeT predict decreasing psychometric functions in bisections of the trained durations, SET predicts the function will also superimpose. There is no relation between green and blue, although both are associated with the 4 s time duration, so SET predicts indifference in choosing between the two colours. The horizontal line in the bottom left panel of Figure 2 demonstrates SET's prediction. LeT subsequently predicts an increasing preference for green over blue as the signal duration increases. The other major differing prediction between SET and LeT is when the two colours associated with the short durations, red and blue, are compared, LeT predicts a U-shaped function while SET predicts a decreasing function similar to the psychometric functions of the other novel colour comparisons (see bottom panels of Figure 2).

The findings of Machado and Keen (1999), Machado and Pata (2005) and Machado and Arantes (2006) are consistent with LeT's predictions. The one

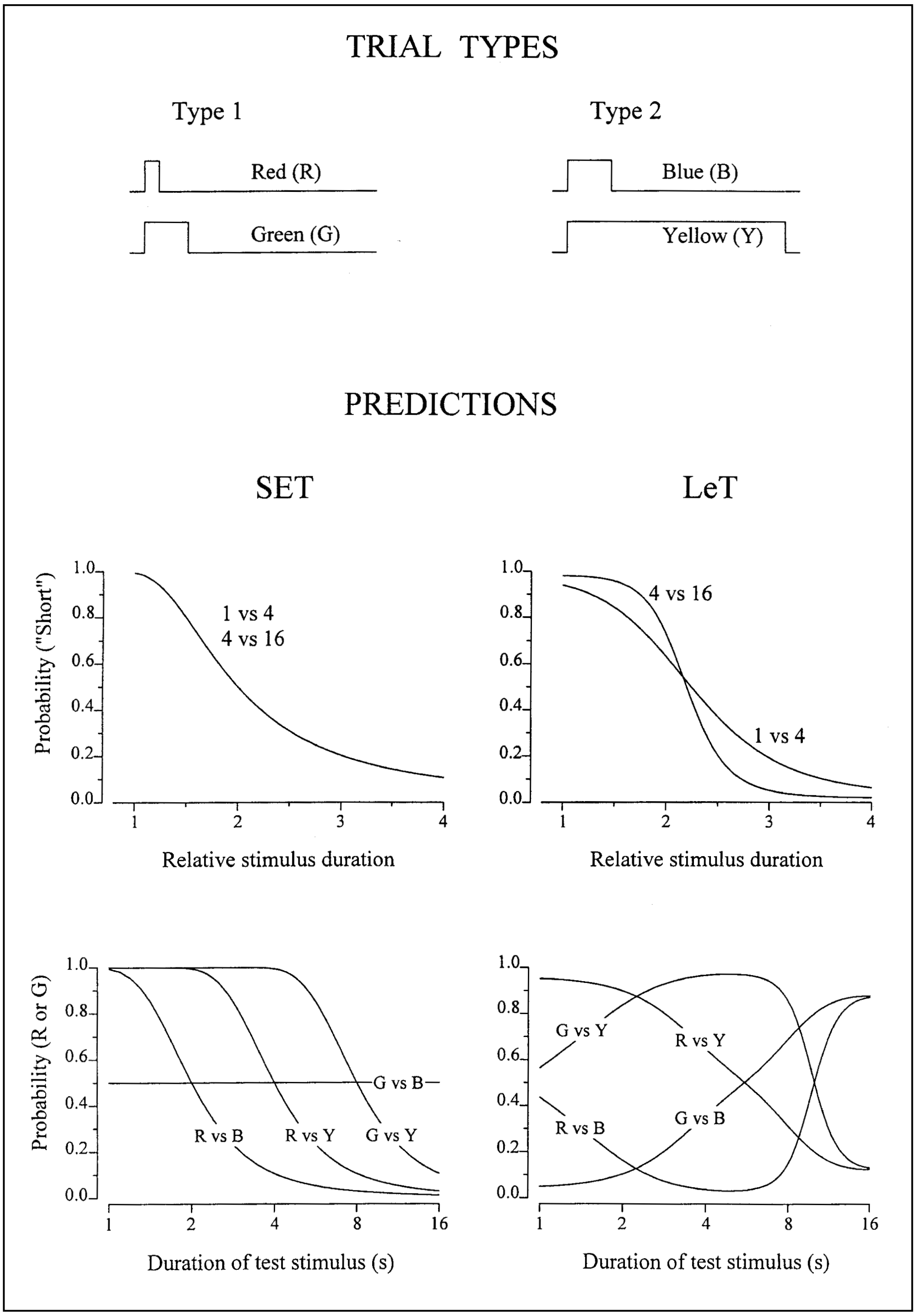


Figure 2. The structure of the double bisection trials (top panel), prediction of SET and LeT with trained durations (middle panel) and novel combinations (bottom panel). From "Learning to Time (LeT) or Scalar Expectancy Theory (SET)? A Critical Test of Two Models of Timing" by A. Machado and R. Keen, 1999, *Psychological Science, 10,* pp. 285-290.

prediction by SET and LeT that Machado and Keen (1999) were unable to demonstrate was in regard to the PSE. In a temporal bisection the PSE is where the probability of choosing either of the two colours are equal. SET predicts the PSE exactly at the geometric mean of the two trained durations, whereas LeT predicts the PSE will be approximately at the geometric mean. The findings of Machado and Keen (1999) showed that of the 16 cases only 5 were within the 95% confidence interval surrounding the geometric mean, 11 cases were outside the interval and had a PSE that was not the geometric mean as SET and LeT predicted. Machado and Keen (1999) stated that neither theory accounted for a PSE considerably below the geometric mean.

Bizo and White (1997) looked at the different predictions made by SET and BeT, and in particular the predictions regarding Weber fractions. Weber fractions are a value that signifies the lease amount of noticeable change. SET argues that a clock and a pacemaker are used to timing behaviours, BeT argues that timing can be explained and accounted for by a pacemaker rate that varies relative to the rate of reinforcement. Bizo and White (1997) found results that are inconsistent with the core assumption of BeT. Their results showed that when Weber Law holds the rate of reinforcement and the pacemaker rate are not precisely proportionally. Bizo and White's (1997) findings support other predictions of BeT as the Weber fractions were not constant after changes in reinforcer density or if reinforcer density was held constant across trial durations, as BeT predicted. In both cases SET incorrectly predicted constant Weber fractions.

SET's double bisection predictions (Figure 2) and superposition of durations with the same S/L (Short/Long) ratio have been shown as incorrect

(Ludvig, Balci, & Longpre, 2008; Machado & Keen, 1999). Ferrara et al. (1997) explain a lack of superposition in previous research by stating that the degree of superposition will vary with the difficulty of the experimental conditions. Trujano and Zamora (2013) argue that experimental conditions, data analysis and individual subjectivity can account for differing findings regarding superposition, they showed this in their study where superposition improved when normalized by the bisection point, but poor superposition if normalized by the shortest duration.

SET is reliant on Weber's Law and generalised Weber's Law, which have been shown to fail by Bizo and White (1997). Despite these critical aspects of SET being shown not to hold in certain timing tasks, it is still the most prominent theory of timing. Different timing theories have been tested with a variety of timing tasks and animals, though a timing task using dogs has not been conducted.

Timing and Dogs

Time is an important stimulus dimension in the life of dogs because it impacts their ability to anticipate future events such as predicting where to be at certain times in order to receive food, or avoiding negative things such as a telling off from their owner. The awareness of a dog's ability to understand time can affect how people approach dog obedience training and how people in general approach animal timing issues.

Examining the performance of dogs on a temporal bisection task with dogs will provide new information on temporal sensitivity in dogs, which can affect how dogs are trained and how humans interact with them. How sensitive dogs might be to the delay of reinforcement might be predicted in part by their performance on a temporal bisection task, as this will show how sensitive they are to the passage of time and if they can discriminate between time intervals that only differ by a few seconds. Temporal bisection results for dogs might differ from those already produced with other animals, since dogs are companion animals as opposed to being regularly used for experimental work. This could affect their results and needs to be considered if comparing their results to those of pigeons or rats who are experimental animals. Dogs are strong in both sensitivity to cueing by humans and training ability when compared to other animals previously used in temporal bisection tasks so this could also affect the results.

There is still much discussion over whether dogs can see in colour or are colour blind. The colour discriminations of red, blue, green and yellow used in standard temporal bisection and double bisection procedures should not pose a problem for dogs as Neitz, Geist and Jacobs (1989) demonstrated that dogs have no difficulty discriminating between separate colours and do not see in black and white. The results from Neitz et al. (1989) demonstrated dogs have dichromatic colour vision, meaning that they can easily distinguish between various shades of blue, green and yellow but cannot distinguish between varying shades of reds and oranges.

The ability of dogs to time is relatively unknown, although Devenport and Devenport (1993) showed that dogs were able to adjust quicker to a changed condition if the new condition followed immediately after the old condition. They also showed that the longer the gap between the two conditions, the worse the performance by the dog, with a 24-hr gap being the longest. This demonstrated that dogs are affected in some sense by time and time delays. Roberts (2002) showed that dogs can track time in terms of the time of day, examples of this are dogs waiting for the mailman at the same time each day and wolves waiting for prey at dusk. Roberts (2002) also stated that dogs could learn to time intervals accurately when there is a specific event such as food occurring after it, which is shown through Pavlov's dog experiments.

Call, Brauer, Kaminski and Tomasello (2003) produced a study involving dogs that looked at how sensitive dogs were to human attention. They tested 12 domestic dogs, 8 females and 4 males that ranged from 1yrs old to 9yrs old, of various breeds. The dogs have no specific training other than general obedience training and were not food deprived throughout the experiment. This study involved two experiments; both began with the experimenter placing a piece of dog food on the floor and saying "Aus" which the dogs were familiar with meaning, don't take it. In Experiment 1 the experimenter than gave a command, either "Geh ab" which the dogs were familiar with meaning, take or repeated "Aus" and left the room or sat on a chair looking straight at the dog. In the second experiment after saying "Aus" the experimenter sat on a chair either with eyes open, eyes closed, doing a distracting activity or with her back turned. Range, Aust, Steurer and Huber (2007) demonstrated that dogs are able to categorise pictures in a visual two-choice procedure where there was no social cueing.

The results showed in Experiment 1 that the dogs took less pieces of food in conditions where they were instructed "Aus" and the experimenter stayed in the room. In Experiment 2 the dogs ate the piece of food significantly less of the time in the eyes open condition, than when the experimenter had their eyes closed, were distracted or had their back turned. These results demonstrate that dogs are aware and sensitive to the attention of humans as they ate less food in conditions where the experimenter could see the dog.

Dogs are regularly used for hunting, herding, general farm work, assisting police, aiding the handicapped and as companion animals. All of these activities require timing and decision making skills. While Call et al. (2003), Range et al. (2007), and Roberts (2002) are a few studies that have involved dogs and demonstrate and tested their ability to make choices based on visual stimuli, timing abilities of dogs remains largely un-researched.

The ability for dogs to discriminate between different time periods has not been looked at and no studies on temporal bisection have been done involving dogs. The present study is a standard temporal bisection procedure and aims to determine whether dogs can discriminate between different time periods, as well as demonstrate the differences, if any, between dogs and other animals previous used in temporal bisection tasks.

Method

Subjects

The subjects were six domestic dogs, three females and three males, of various breeds. Aged between 2 and 6-years old. They were all experimentally naïve; water was available during and after each training session. The owners of the dogs were asked to not give breakfast to the dogs on days when they participated in the experiment. The Animal Ethics Committee for the University of Waikato gave permission for the dogs to be used as part of this study (protocol number: 855).

Apparatus

A box with three vertical levers was used in this experiment (57.5L x 22.5W x 15H cm) the top panel is made of transparent acrylic while the sides and floor is made of wood and painted white. The three vertical levers (10 x 1 cm) stick straight up from the box. Each side could light up red or blue as shown in Figure 3 (middle panel), and the whole top panel would light up green at the start of each trial (Figure 3, top panel). A food dispenser (Manners Minder) dispensed 3-5 dog-food pellets onto a food tray after a correct response. A laptop controlled the experiment and recorded the data.

Procedure

Pretraining. Each dog was given one session of Manners Minder training**.** A manners minder is a food dispensing system for dry dog food (Figure 3, bottom panel). The box and levers were not presented and food reinforcement of dog food

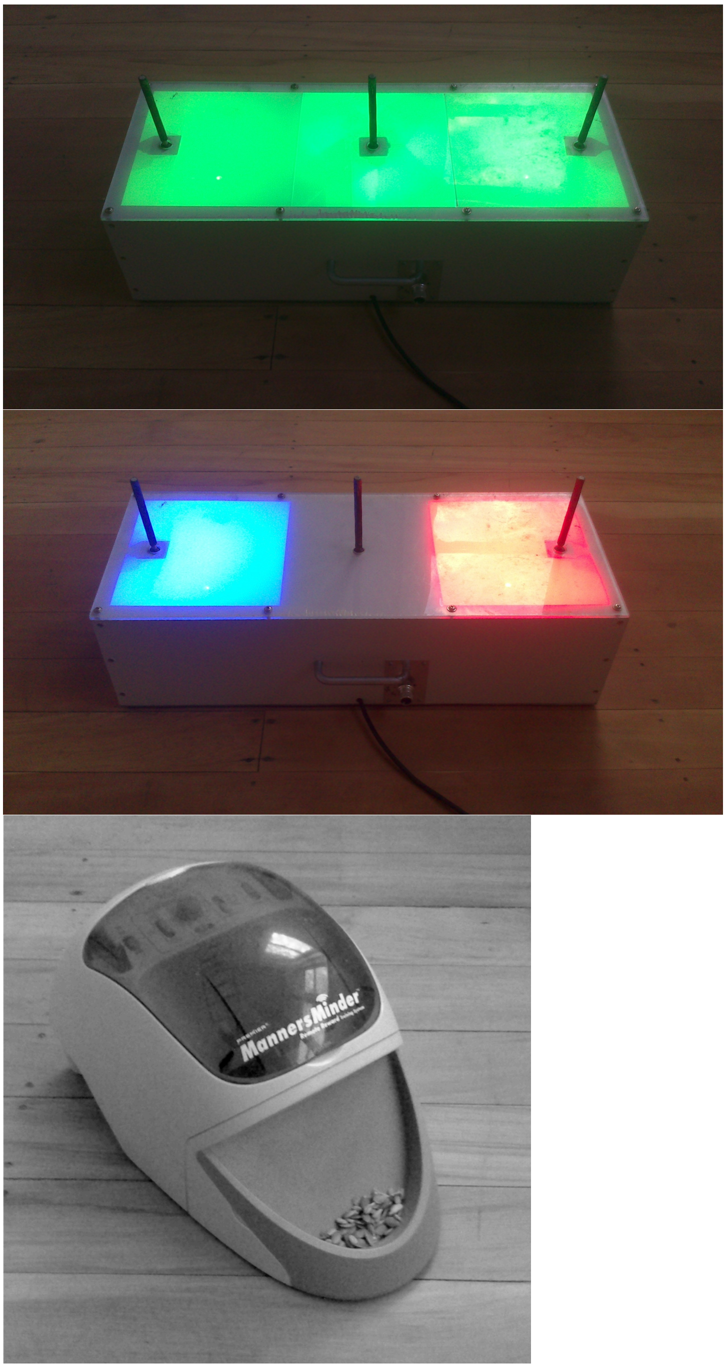


Figure 3. Photos of the Apparatus used in the experiment. Green light during signal duration (top panel), red and blue lights for the response options (middle panel) and dog-food reinforcement delivery system, manners minder (bottom panel).

pellets were delivered every 30 s for 30 minutes by the manners minder. In the next session the dogs were trained to push the vertical levers on the box. Ten responses on the middle lever were reinforced, ten responses on the left lever were reinforced, and then ten responses on the right lever were reinforced. The dogs were trained to press the middle lever to start each trial, which started a green light of 1.0 s duration then a blue light lit up either the left or right lever and reinforcement was delivered if the lit lever was pressed. When the animal made the correct response 70% of a daily session the condition changed and the green light was a 4.0 s signal and only the red response key lit up. After 70% correct responding on only 4.0 s red trials, alternating trials of only blue and only red were conducted for five successive sessions.

Training. In training the dogs pressed the middle lever to begin each trial and either a 1.0 s or 4.0 s green light was shown, after the signal duration both a red light and a blue light came on, on either the left or right side and the middle lever was left unlit. After a response was made, by pressing either of the side levers (pressed the middle lever had no effect) all lights went off. If the response was correct reinforcement was immediately given by the manners minder. If the response was incorrect no reinforcement was given. The dog pressed the middle lever to start the next trial. An experimental session lasted for 50 trials. A record of correct and incorrect responses was kept for blue and red responses following the two signal durations. Two concessive daily sessions of 80% correct or higher ended the training period.

Testing. When dogs met the accuracy criteria for ending the training sessions, they experienced one testing session. The same conditions as training were used for testing except five intermediate durations were presented as well as the two extreme signal durations of 1.0 and 4.0 s. Correct responses after the two extreme signal durations were still reinforced. Responses after the five intermediate durations received no reinforcement. The intermediate durations presented were logarithmic intervals of the two original signal durations. On each trial there were 7 possible durations for the green light, the two trained durations and the 5 intermediate logarithmic durations between them. The three middle durations are at the harmonic, geometric and arithmetic mean of the two trained signal durations. In each trial the probability of a trained extreme duration being presented was 0.25 for each of the two trained durations. So reinforcement was available on half the trials in a testing session, and the other half were intermediate durations.

The same training and testing procedure was followed for the other three conditions using 0.5 and 2.0 s, 2.0 and 8.0 s, and 4.0 and 16.0 s as the extreme duration pairs during training.

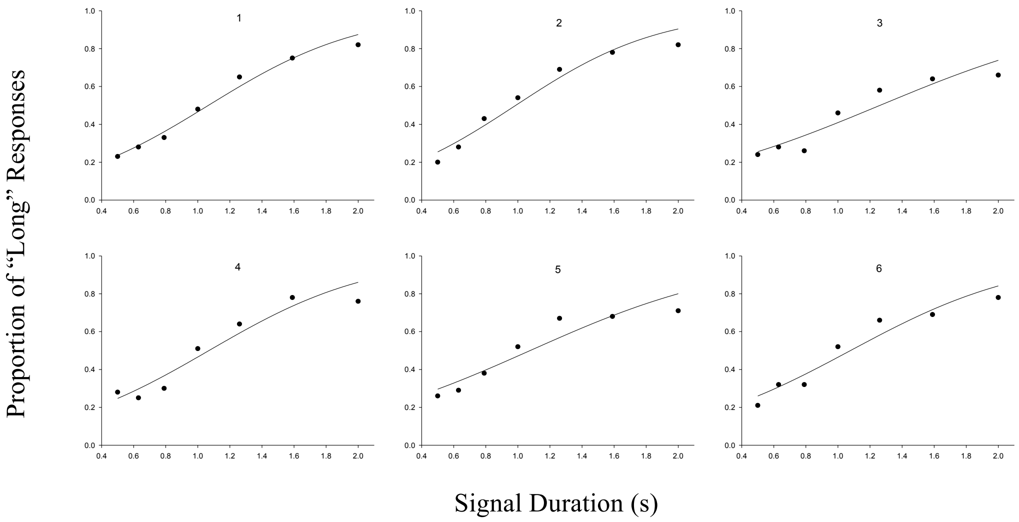
Results

In the testing condition there were seven signal durations presented, the two extreme trained durations and five intermediate durations. The intermediate durations used in testing for each duration range were calculated by multiplying the smallest value by a constant 1.26, this ensured the proportion change was constant. This also resulted in all intermediate durations being spaced equal logarithmic intervals apart. Every dog was exposed to the harmonic, geometric and arithmetic means of the trained durations as each mean was an intermediate signal duration in the testing of each condition.

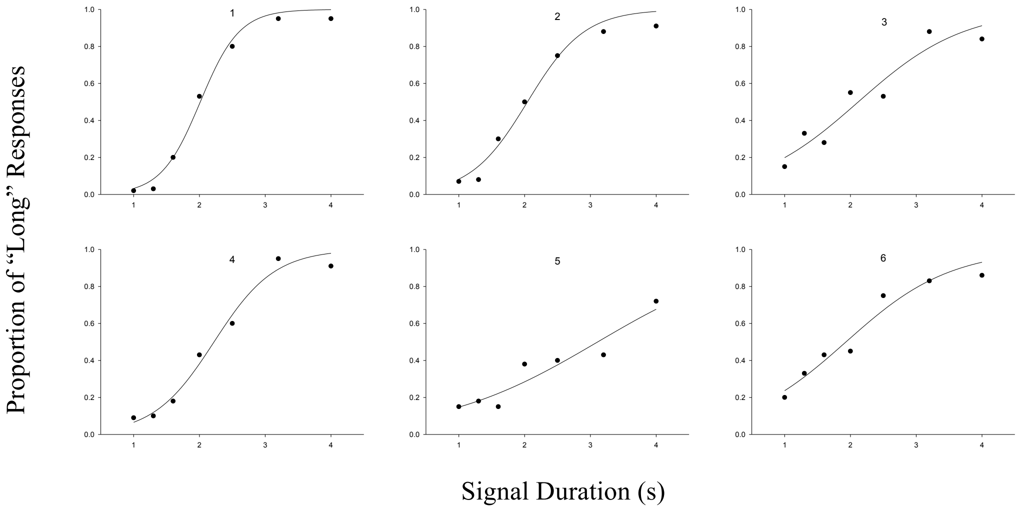
As predicted the number of "Long" responses was low at the shorter signal durations and progresses higher as the signal duration increases for all subjects across all four conditions. The model (Equation 1) was fitted to the individual responses proportions for each of the dogs under all four conditions using non-linear least squares regression and the best fits are presented as smooth lines in Figure 4. Performance was similar across subjects and conditions, with the exception of condition 1.0 – 4.0 s, where the psychometric function that described subject 5's pattern of responding was much flatter compared to the other subjects. While the 2.0 – 8.0 s condition produced responding from all six subjects that was very similar. Figure 4 also highlighted differences between conditions, illustrating that individual accuracy was consistently lower for the 0.5 – 2.0 s and 4.0 – 16.0 s conditions.

The model tends to overestimate performance at the longer durations of each condition but overall fit well for all of the subjects across all four conditions. The proportion of long responses across all the subjects was averaged for each of

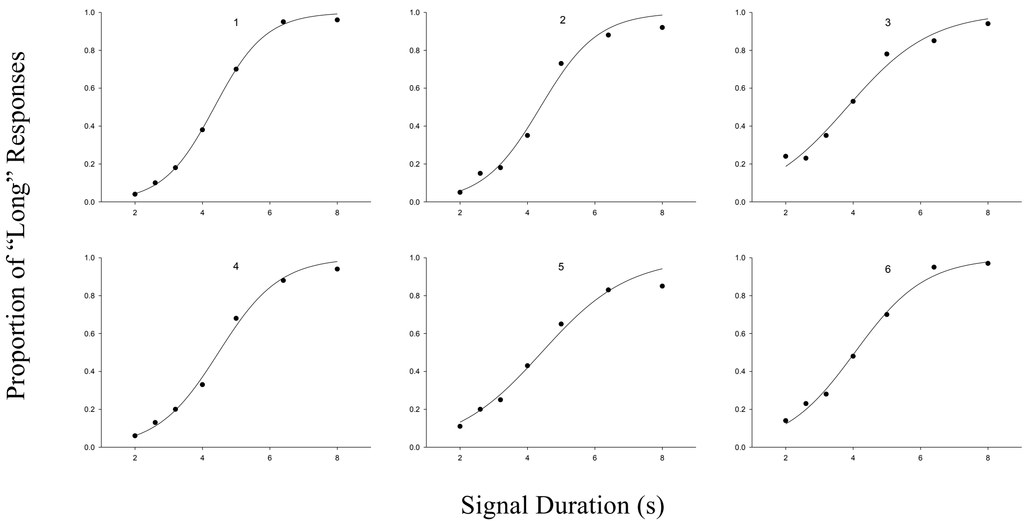
Condition 0.5 – 2.0 s



Condition 1.0 – 4.0 s



Condition 2.0 – 8.0s



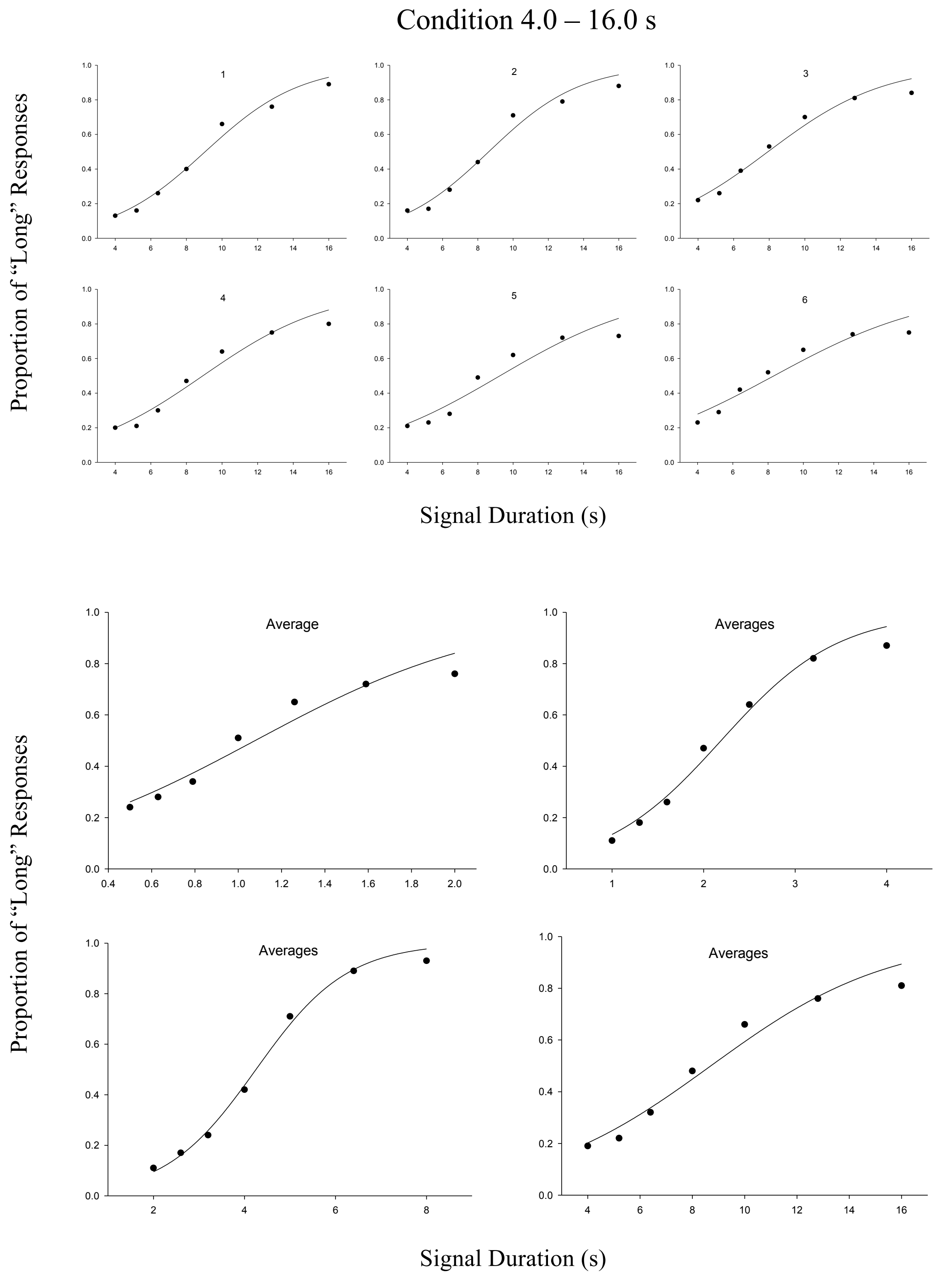


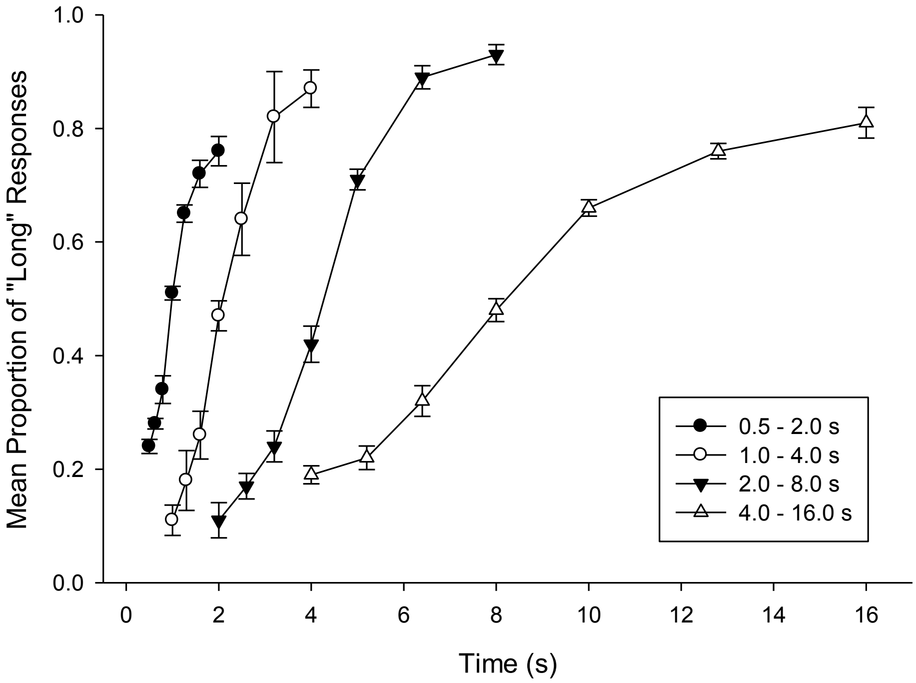
Figure 4. Proportion of "Long" responses plotted as a function of signal duration. The smooth lines through the data points were fitted by Equation 1.

the four signal duration ranges and Equation 1 was then fitted to the averaged data, and the best fits represented as smooth lines through the data in Figure 4. The functions are steeper for the 1.0 – 4.0 s and 2.0 – 8.0 s conditions. The fits of the model to the data were also better for these conditions.

The average proportion of long responses across all the subjects was calculated for each of the four signal duration ranges and a mean psychometric function was created. The Y axis in both panels of Figure 5 shows the proportion of “Long” responses. Under all four conditions the proportion of “Long” responses increased with the signal duration, produced similar patterns of responding for each condition, to varying degrees of accuracy. This pattern is consistent with other findings from a temporal bisection procedure.

Figure 5 (bottom panel) shows the averages of each of the four conditions when scaled in logarithmic units, this means that responding can be compared across the relative durations in all four conditions. The 0.5 – 2.0 s condition is the shortest duration range and has the smallest short extreme duration of 0.5 s. This condition produced the flattest function, indicating that subjects found discriminating between the durations within this condition the hardest. The 2.0 – 8.0 s condition produced the steepest function, indicating discriminability was high. All four conditions visually demonstrated a 0.5 probability of responding “Long” around or just past the 4^{th} signal duration, which corresponds to the geometric mean of the two trained durations in each condition.

Figure 5 (top panel) shows the mean proportion of long responses as a function of time. The slopes of the functions flatten as time increased even though the S/L (Short/Long) ratio remains constant at 1:4, demonstrating that



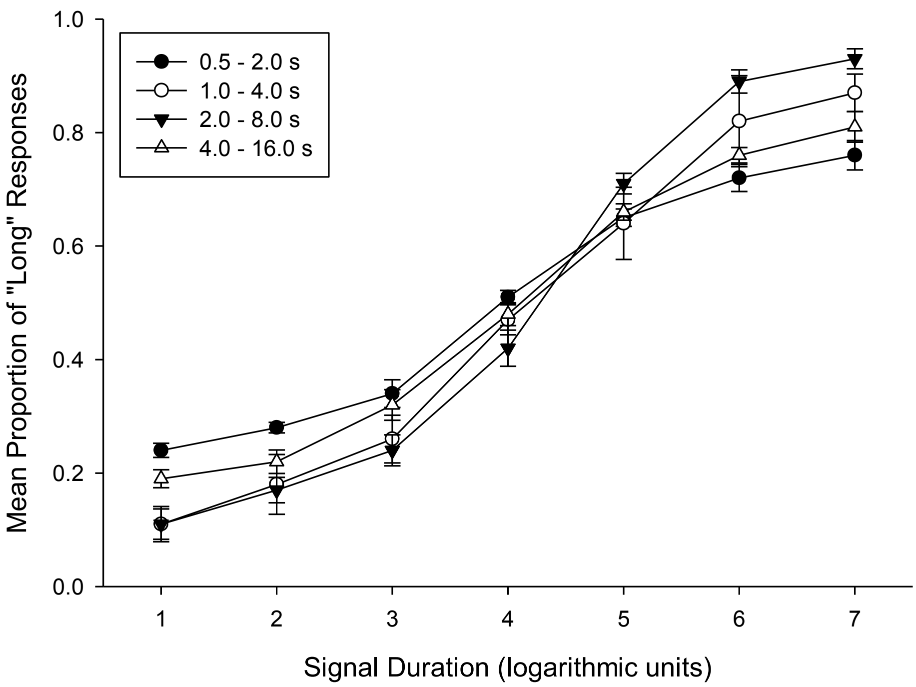


Figure 5. Proportion of long responses as a function of time (top panel), as a function of signal duration in logarithmic units (bottom panel). The four duration ranges used were 0.5 – 2.0, 1.0 – 4.0, 2.0 – 8.0 and 4.0 – 16.0 s.

superposition did not occur. The two middle conditions 1.0 – 4.0 s and 2.0 – 8.0 s are shown to be in about the same accuracy range. In the first condition the dogs were required to press the lever when the light was blue after a 1.0 s signal and red after a 4.0 s signal. It took the dogs 20.2 days on average to meet the criteria of 80% correct over two successive sessions, and 8.5 days for the 0.5 – 2.0 s discrimination, 8.7 days for the 2.0 – 8.0 s discrimination and 10.3 days for the 4.0 – 16.0 s discrimination. There was also a difference between subjects within the same condition as to how many training sessions were required before the criteria to begin the testing condition was met. The largest different between two subjects on the same condition was 14 days.

The PSE is the signal duration that when presented has a 0.5 probability of eliciting a "Long" response from the subjects. There are various methods of calculating the PSE that all produce similar results. Here the PSE was calculated by the individual interpolation (IINT) method, where the PSE is interpolated from the psychometric function of the individual subjects and these values were then averaged across dogs. Based on this method the PSE for the 0.5 – 2.0 s interval was 1.0 s, this is also the geometric mean of this condition. The PSE for the next three conditions were 2.28 s (1.0 – 4.0 s), 4.26 s (2.0 – 8.0 s) and 8.23 s (4.0 – 16.0 s), all extremely close to the geometric means of 2.0 s, 4.0 s and 8.0 s, respectively.

Figure 6 shows the PSE as a function of the geometric mean of the two extreme trained durations. How close the data points are to the geometric mean line (diagonal line, Figure 6) indicate that the geometric mean and the PSE are extremely close, and for the first condition overlap.

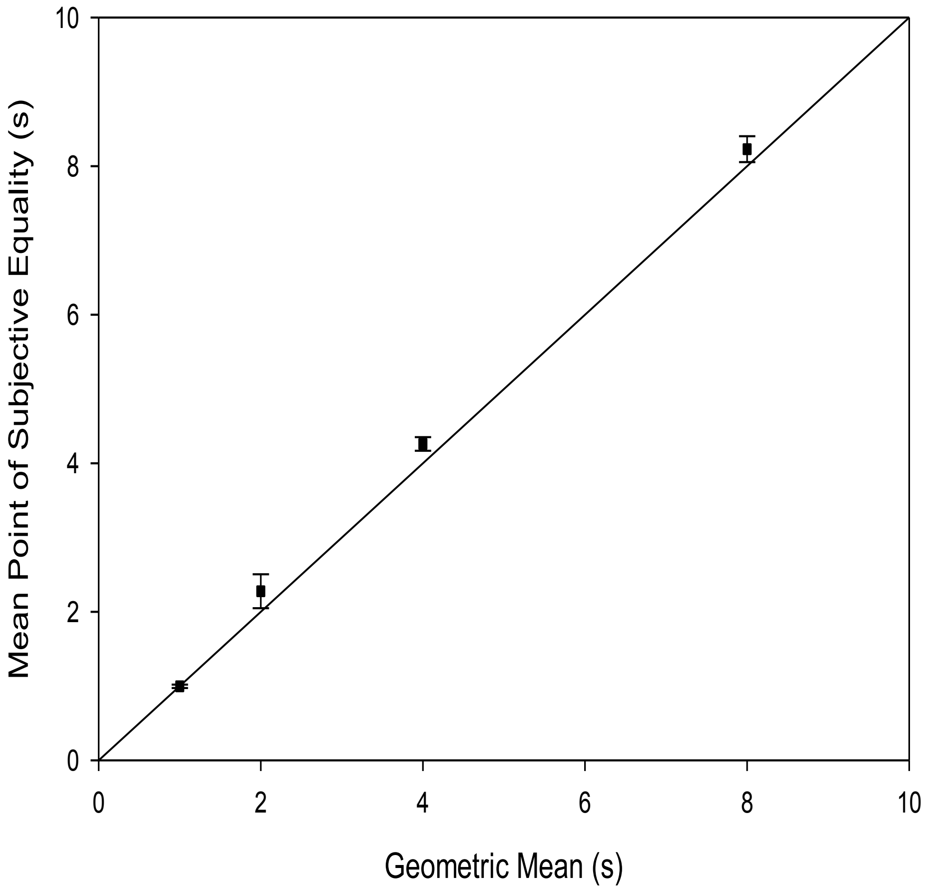


Figure 6. The Point of Subjective Equality as a function of the geometric mean of the two trained extreme durations in each condition.

The PSE was closer to the geometric mean than the arithmetic or harmonic mean for all four conditions and individual testing was a clear indicator of this, as this was the case for every dog in all four conditions except for one dog, subject 5, in the 1.0 – 4.0 s condition where the PSE was closer to the arithmetic mean. The variance accounted for ranged from 93%-99%, with an average of 95.7% across all conditions. The Weber fractions were calculated by dividing the standard deviation by the mean to give a measure of discrimination accuracy.

As shown in Figure 7 the Weber fractions were largest for the shortest condition, which was 0.5 – 2.0 s (0.935), smaller for the 1.0 – 4.0 s and 2.0 – 8.0 s conditions 0.529 and 0.438 respectively and increased again for the longest 4.0 – 16.0 s condition that had a Weber fraction of 0.730. This created a U-shaped function. These Weber fractions indicated that the signal ranges of 1.0 – 4.0 s and 2.0 – 8.0 s were more accurately discriminated than the others. Table 1 shows the parameter values from each dog under each condition, including the mean and standard deviation used to calculate the Weber fraction.

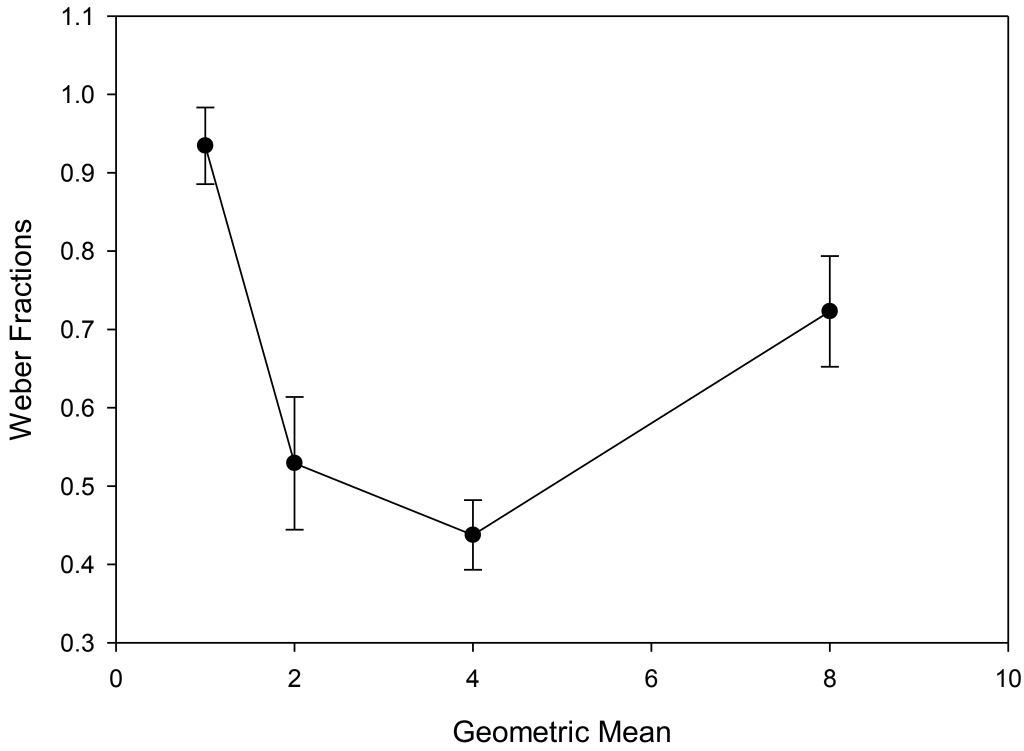


Figure 7. Weber fraction as a function of the geometric mean across the four conditions.

Table 1. *The mean, standard deviation (Sd), Weber Fractions (WF), R-squared, and Standard error of estimate (SEE) for all subjects, for each of the four conditions.*

Subjects	Mean	Sd	WF	Rsqr	SEE
0.5 – 2.0 s					
1	1.065	0.875	0.821	0.980	0.037
2	0.986	0.822	0.833	0.956	0.056
3	1.261	1.295	1.027	0.890	0.067
4	1.069	0.929	0.870	0.924	0.070
5	1.078	1.004	0.932	0.934	0.062
6	1.076	1.210	1.124	0.889	0.070
Average	1.089	1.023	0.935	0.929	0.060
1.0 – 4.0 s					
1	2.013	0.542	0.269	0.991	0.044
2	2.043	0.789	0.386	0.979	0.058
3	2.119	1.461	0.690	0.917	0.088
4	2.222	0.838	0.377	0.983	0.053
5	1.932	1.446	0.749	0.954	0.062
6	3.110	2.188	0.704	0.904	0.070
Average	2.240	1.211	0.529	0.955	0.063
2.0 – 8.0 s					
1	4.355	1.383	0.317	0.999	0.016
2	4.384	1.547	0.353	0.985	0.050
3	3.827	2.260	0.591	0.979	0.047
4	4.478	1.646	0.367	0.993	0.033
5	4.050	1.897	0.468	0.994	0.029
6	4.418	2.336	0.529	0.980	0.048
Average	4.252	1.845	0.438	0.988	0.037

4.0 – 16.0 s					
1	9.049	4.869	0.538	0.982	0.044
2	8.617	4.730	0.549	0.973	0.054
3	7.927	5.933	0.749	0.971	0.048
4	8.919	6.433	0.721	0.961	0.055
5	8.329	8.285	0.995	0.929	0.061
6	9.249	7.658	0.828	0.916	0.073
Average	8.682	6.318	0.73	0.955	0.054

Discussion

Temporal Bisection

Across conditions, each subject had a lower proportion of “Long” responses at the shorter signal durations and progressed higher as the signal duration increased. Although the findings from this study demonstrated the varying degrees of discrimination across durations, with the smallest and largest ranges producing much flatter discriminations.

As with most temporal bisection procedures the harmonic, geometric and arithmetic means were included as intermediate durations in the testing phase for each condition. This provided a clear indication of which mean was closest to the PSE. The geometric mean was the closest to the PSE in all conditions. This replicates the findings of other temporal bisection procedures done using rats (Church & Deluty, 1977; Meck, 1983; Raslear, 1983), pigeons (Stubbs, 1976) and mice (Penney et al., 2008) where the PSE was also approximately at the geometric mean. This study was conducted with dogs, a species that was not involved in any previous temporal bisection studies and these findings add support to the general conclusion that animals bisect at the geometric mean.

The 0.5 – 2.0 s signal range was the shortest duration range tested in this study and produced the flattest function, indicating that subjects were less accurate when discriminating between the durations within this condition. This result found in the present study supports the previous findings of Zelanti and Droit-Volet (2011), who found the accuracy of responding on their shortest signal range (0.5 – 1.0 s) significantly lower than their other three conditions (1.25 – 2.5 s, 4.0 – 8.0 s, 15.0 – 30.0 s). Further research should look closely at durations less than 1.0 s, as smaller durations offer different findings and might be able to provide more evidence for or against current timing theories and help rule out any incorrect assumptions made.

The results of the present study are consistent with the majority of results found by Church and Deluty (1977), although Church and Deluty (1977) did not test a signal duration less than 1.0 s to compare to. Figure 8 shows their averaged data expressed as proportion of "Long" responses as a function of time. Church and Deluty (1977) conducted temporal bisection tasks with four signal duration pairs, 1.0 – 4.0, 2.0 – 8.0, 3.0 – 12.0, and 4.0 – 16.0 s. The smooth curves are fits of Equation 1 to the data. The fits showed in Figure 8 are similar to results found in the present study, with Equation 1 overestimating performance at the longer durations of each condition. Table 2 presents the parameter values from fitting Equation 1 to Church and Deluty's data. The model generally provides a good fit to the data through all four conditions with the variance accounted for ranging from 97%-99%, as shown in Table 2. This demonstrated that Equation 1 is a valuable model when fitting to temporal bisection data.

Weber's Law

Weber's Law states that the Weber fractions will remain consistent when given a constant S/L (Short/Long) ratio. This study had a constant S/L ratio of 1:4 through out all four conditions. The results showed that as stimulus durations get longer the Weber fractions get smaller and then increase again when the stimulus duration is very long. Figure 7 shows the Weber fractions as a function of the geometric means for the four conditions and is clearly a U-shaped function. These

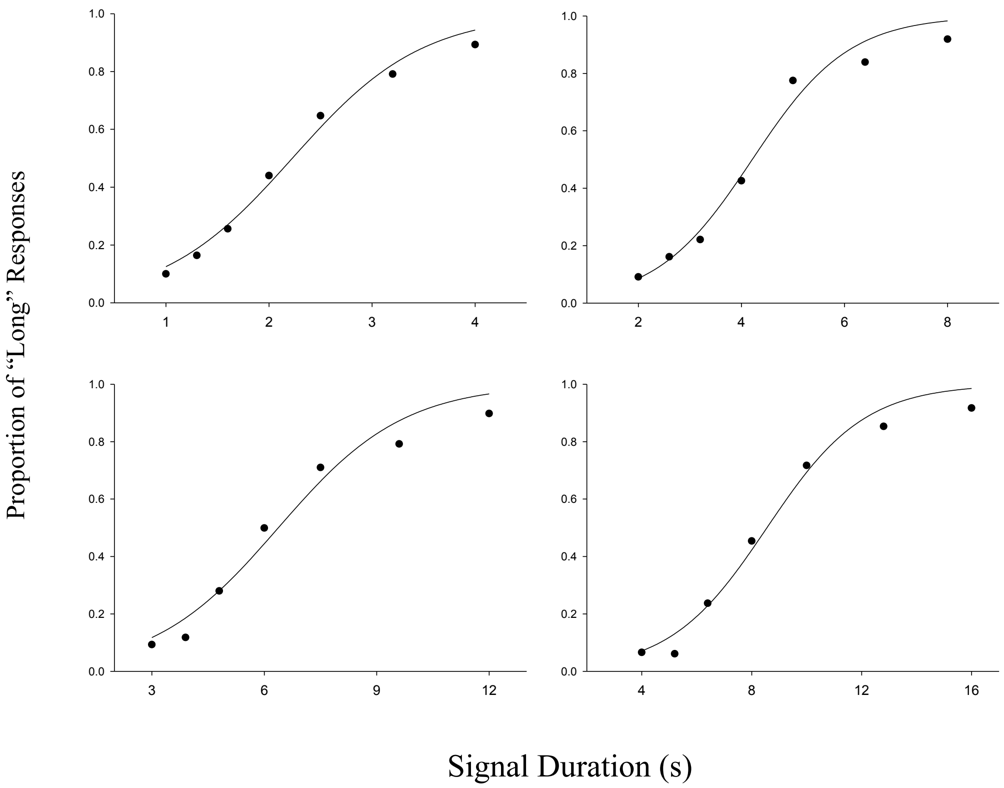


Figure 8. The data was taken from Church and Deluty (1977). The four duration ranges they used were 1.0 – 4.0, 2.0 – 8.0, 3.0 – 12.0 and 4.0 – 16.0 s, and shows the proportion of "Long" responses plotted as a function of signal duration. The smooth lines through the data points were fitted by Equation 1.

Table 2. *The mean, standard deviation (Sd) and R-squared values from Church and Deluty's (1977) averaged data under each of the four conditions when Equation 1 was fitted to the data.*

Condition	Mean	Sd	Rsqr
1.0 – 4.0 s	2.226	0.054	0.987
2.0 – 8.0 s	4.211	0.134	0.977
3.0 – 12.0 s	6.374	0.254	0.966
4.0 – 16.0 s	8.554	0.252	0.981

findings are inconsistent with Weber's Law and its generalized form, which predict a flat function for Weber's Law or a reverse J-shaped function of the generalized Weber's Law (Getty, 1975).

These findings are consistent, however, with reports from a growing number of research papers such as Bizo et al. (2006), Zeiler (1991), and Lavioe and Grondin (2004) that have demonstrated the failure of Weber's Law to predict substantial increases in the standard deviation of longer duration ranges in a temporal bisection procedure. This study also adds support to the idea of a U-shaped function being a more accurate description of Weber fractions as demonstrated by Bizo et al. (2006).

It is widely accepted that short durations results in high Weber fractions and generalized Weber's Law is believed to account for that, further research should look at the longer durations in more detail and attempt to replicate this finding of a U-shaped function, particularly in the longer duration ranges to see if this finding holds up.

Theories of Timing

Figure 5 (top panel) demonstrated an inconsistency with scalar property, which SET relies on. All four of the conditions have the same S/L ratio of 1:4 but when the x-axis was normalized the data points did not superimpose, indicating that they did not align on the y-axis and demonstrating a failure of scalar property. This finding is unlike previous research, which has been able to demonstrate scalar property (Droit-Volet, 2002; Droit-Volet, Clement & Wearden, 2001), Machado (1997) demonstrated superposition of FI performance and concluded a Weber-like property to their temporal bisection procedure.

Results have provided support for LeT's prediction (Figure 2) that when durations with a constant S/L ratio are compared the shorter duration range will produce a flatter function (Machado & Keen, 1999). Figure 5 showed this with the short 0.5 – 2.0 s condition being flatter than the other conditions.

These findings have added support to LeT's prediction about not producing superposition and BeT's predictions that Weber fractions are not constant and produced U-shaped functions, not flat or J-shaped. This concurs with previous findings from Machado and Keen (1999), Bizo et al. (2006), Bizo and White (1997), Machado and Arantes (2006) and Ludvig et al. (2008) who all found results that's either demonstrated a substantial increase in Weber fractions at long durations or a lack of superposition on Short/Long discriminations of equal ratios. LeT and BeT do not rely on Weber Law as part of their core assumptions so are not negatively affected when it is shown to have failed. LeT even predicts deviations and failures of Weber Law under certain conditions due to learned associations (Machado & Keen, 1999), so a failure of Weber's Law can support LeT's prediction.

SET, however, is reliant on both Scalar timing and Weber's Law as part of the theories core assumptions. The only finding that provided support for SET was the PSE being at the geometric mean in the 0.5 – 2.0 s condition. Under the other three conditions LeT's predictions were more accurate, predicting the PSE at approximately the geometric mean (see Figure 6) rather than the harmonic or arithmetic means.

In the present study the intermediate durations were spaced logarithmically between the short and long training stimuli. Wearden et al. (1997) found that by spacing the intermediate durations linearly the PSE was reduced.

Further research could reproduce the present study with linearly spaced intermediate durations and see whether the PSE was reduced and if it was then at the geometric mean, which would be consistent with SET's predictions.

Failure of predictions and fundamental aspects of SET has been shown in the results and through previous research. Further research into the different predictions made by LeT and BeT is needed, and direct comparisons between the two theories would benefit the field. It is essential that research that supports the predictions and core assumptions made under these theories is investigated to determine if either of these two theories are complete and can explain timing or if further development of new and existing theories is needed.

Timing and Dogs

The temporal bisection results demonstrated new findings in the area of timing with dogs and indicate the ability of dogs to discriminate between time durations. The results were consistent with those of other animal studies, PSE at the geometric mean and U-shaped Weber fractions. Further research into temporal bisection with dogs is needed to ensure that certain findings were not due to experimental conditions.

Conducting research with companion animals rather than laboratory animals did not seem to effect the results as after the initial training the dogs were familiar with the general experimental conditions, although the size and mobility of the subjects and the set up of the response system may have had an impact on results. The response levers were very close to the start lever for each trial (See Figure 3) and during the 0.5 – 2.0 s condition their appeared to be occasions when the response lever was hit accidently immediately after the shorter durations were presented. This could help explain the discrepancy between this condition and the other three. There was also a lot of adjunctive behaviour occurring around the response box while the signal duration was being presented which may have affected the results. Having a responding system where the two possible responses required a more direct physical effort, for example having the two response options presented at opposite sides of a room, would be beneficial in ensuring this does not affect results.

The ability and extent that dogs can discriminate between time durations is important because the findings influence how dog training is approached. Successful timing studies with dogs will enable faster and more accurate training in all capacities, from domestic obedience training, more advanced hunting and retrieving skills, to extreme training such as specialized police work and service dogs. Further research into the other abilities of dogs will provide the knowledge to create training programs that are more successful, with a faster mastery of skills, less incorrect responses and less risk of surprise behaviour.

The findings of this study demonstrated some dogs were better at this task and learned the discrimination faster. In some applied settings individual differences in rates of learning are important, such as training service dogs, with dogs that are slow learners costing more to train and often never graduating from training to service. Further research and development on the reasons for differences in rates of learning and performance across individual animals could help produce skill tests across a range of stimulus dimensions and tasks that might identify which dogs might succeed and survive service dog training. This would be beneficial to society as not only will the training be of a higher quality but the dropout rate reported by dog training centres of at least 50% (Weiss, 2002) would decrease, saving time, money and resources in the future. This would expand the areas where dogs are used to assist people in making life easier or preforming tasks that require specific abilities such as an advanced sense of smell.

In conclusion results from this study support general findings about how animals perform in a temporal bisection task, but offer new findings for durations less than 1.0 s. Further research should look at explaining the findings during the shortest condition (0.5 – 2.0 s), to determine if experimental conditions, mainly the response apparatus, can account for the difference found between this condition and the three others. As well as test whether logarithmically spaced compared to linearly spaced makes any difference.

The cognitive abilities of dogs remains a largely untouched research area and further research in to their abilities, limitations and successful training procedures is necessary to expand our knowledge and utilize the potential of dogs as assets within our communities and general society.

References

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

Balsam, P., Sanchez-Castillo, H., Taylor, K., Van Volkinburg, H., & Ward, R. D. (2009). Timing and anticipation: Conceptual and methodological approaches. *European Journal of Neuroscience, 30*, 1749-1755.

Bizo, L. A., & White, K. G. (1997). Timing with controlled reinforcer density: Implications for models of timing. *Journal of Experimental Psychology: Animal Behavior Processes, 23,* 44–55.

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*, 201-210.

Call, J., Brauer, J., Kaminski, J., & Tomasello, M. (2003). Domestic dogs (*canis familiaris*) are sensitive to the attentional state of humans. *Journal of Comparative Psychology, 117*, 257-263.

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

Devenport, J. A., & Devenport, L. D. (1993). Time-dependent decisions in dogs (*Canis familiaris*). *Journal of Comparative Psychology, 107,* 169-173.

Droit-Volet, S. (2002). Scalar timing in temporal generalization in children with short and long stimulus durations. *Quarterly Journal of Experimental*

Psychology A, 55, 1193–1209.

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

Droit-Volet, S., Clement, A., & Wearden, J. (2001). Temporal generalization in 3- to 8- year-old children. *Journal of Experimental Child Psychology, 80,* 271–288.

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

Getty, D. J. (1975). Discrimination of short temporal intervals: A comparison of two models. *Perception and Psychophysics, 18,* 1-8.

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

Gibbon, J. (1991). Origins of scalar timing. *Learning and Motivation, 22,* 3-38.

Gibbon, J., & Church, R. M. (1984). Sources of variance in an information processing theory of timing. In H. L. Roitblat., T. G. Bever., & H. S. Terrace (Eds.), *Animal Cognition* (pp. 465-488). NJ: Hillsdale.

Killeen, P. R., & Fetterman, J. G. (1988). A behavioral theory of timing. *Psychological Review, 95,* 274-295.

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

Lavoie, P., & Grondin, S. (2004). Information processing limitations as revealed by temporal discrimination. *Brain and Cognition, 54*, 198-200.

Lee, K. H., Bhaker, R. S., Mysore, A., Parks, R. W., Birkett, P. B. L. and Woodruff, P. W. R. (2009). Time perception and its neuropsychological correlates in patients with schizophrenia and in healthy volunteers. *Psychiatry Research, 166,* 174-183.

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

Ludvig, E. A., Balci, F., & Longpre, K. M. (2008). Timescale dependence in a conditional temporal discrimination procedure. *Behavioural Processes, 77,* 357-363.

Machado, A. (1997). Learning the temporal dynamics of behavior. *Psychological Review, 104,* 241–265.

Machado, A., & Arantes, J. (2006). Further tests of the scalar expectancy theory (SET) and the learning-to-time (LeT) model in a temporal bisection task. *Behavioural Processes, 72,* 195–206.

Machado, A., & Keen, R. (1999). Learning to Time (LET) or Scalar Expectancy Theory (SET)? A critical test of two models of timing. *Psychological Science, 10,* 285–290.

Machado, A., & Pata, P. (2005). Testing the scalar expectancy theory (SET) and the learning to time model (LeT) in a double bisection task. *Learning & Behavior, 33,* 111–122.

Malapani, C., & Fairhurst, S. (2002). Scalar timing in animals and humans. *Learning and Motivation, 33,* 156-176.

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

Neitz, J., Geist, T., & Jacobs, G. H. (1989). Color vision in the dog. *Visual Neuroscience, 3,* 119-125.

Penney, T. B., Gibbon, J., & Meck, W. H. (2008). Categorical scaling of duration bisection in pigeons (Columba livia), mice (Mus musculus), and humans (Homo sapiens). *Psychological Science, 19,* 1103–1109.

Raslear, T. G. (1983). A test of the Pfanzagl bisection model in rats. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 49–62.

Roberts, W. A. (2002). Are animals stuck in time? *Psychological Bulletin, 123,* 473-489.

Range, F., Aust, U., Steurer, M., & Huber, L. (2008). Visual categorization of natural stimuli by domestic dogs. *Animal Cognition, 11,* 339-347.

Siegel, S. F., & Church, R. M. (1984). The decision rule in temporal bisection. *Annals New York Academy of Sciences, 423,* 643–645.

Stubbs, A. (1976), Scaling of stimulus duration by pigeons. *Journal of the Experimental Analysis of Behavior, 26,* 15-25.

Taylor, P. E., Haskell, M., Appleby, C., & Waran, N. K. (2002). Perception of time duration by domestic hens. *Applied Animal Behaviour Science, 76,* 41-51.

Trujano, R. E., & Zamora, O. (2013). Double bisection of auditory temporal intervals by humans. *Psychological Research, 77,* 463-479.

Wearden, J. H., & Ferrara, A. (1995). Stimulus spacing effects in temporal bisection by humans. *Quarterly Journal of Experimental Psychology B, 48,* 289–310.

Wearden, J. H., Rogers, P., & Thomas, R. (1997). Temporal bisection in humans with longer stimulus durations. *Quarterly Journal of Experimental Psychology B, 50,* 79–94.

Weiss, E. (2002). Selecting shelter dogs for service dog training. *Journal of Applied Animal Welfare Science, 50,* 43-62.

Zeiler, M. D. (1991). Ecological influences on timing. *Journal of Experimental Psychology: Animal Behavior Processes, 17*, 13-25.

Zelanti, P. S., & Droit-Volet, S. (2011). Cognitive abilities explaining age-related changes in time perception of short and long durations. *Journal of Experimental Child Psychology, 109,* 143-157.

Appendix A

A: Excel files with raw data are attached on accompanying CD.

Appendix B

B: Excel file with the summary data is attached on accompanying CD.

Appendix C

C: Ethics application and approval for this study is attached on accompanying CD.

Appendix D

D: Sigmaplot notebooks with the raw figures are attached on accompanying CD.